METHOD OF SEQUENCING A NUCLEIC ACID

RELATED APPLICATIONS

This application claims priority to USSN 09/664,197, filed September 18, 2000 and USSN 09/398,833, filed September 16, 1999. The contents of these applications are incorporated herein by reference in their entireties.

FIELD OF THE INVENTION

The invention relates to methods and apparatuses for determining the sequence of a nucleic acid.

BACKGROUND OF THE INVENTION

Many diseases are associated with particular DNA sequences. The DNA sequences are often referred to as DNA sequence polymorphisms to indicate that the DNA sequence associated with a diseased state differs from the corresponding DNA sequence in non-afflicted individuals. DNA sequence polymorphisms can include, *e.g.*, insertions, deletions, or substitutions of nucleotides in one sequence relative to a second sequence. An example of a particular DNA sequence polymorphism is 5'-ATCG-3', relative to the sequence 5'-ATGG-3'at a particular location in the human genome. The first nucleotide 'G' in the latter sequence has been replaced by the nucleotide 'C' in the former sequence. The former sequence is associated with a particular disease state, whereas the latter sequence is found in individuals not suffering from the disease. Thus, the presence of the nucleotide sequence '5-ATCG-3' indicates the individual has the particular disease. This particular type of sequence polymorphism is known as a single-nucleotide polymorphism, or SNP, because the sequence difference is due to a change in one nucleotide.

Techniques which enable the rapid detection of as little as a single DNA base change are therefore important methodologies for use in genetic analysis. Because the size of the human genome is large, on the order of 3 billion base pairs, techniques for identifying polymorphisms must be sensitive enough to specifically identify the sequence containing the polymorphism in a potentially large population of nucleic acids.

Typically a DNA sequence polymorphism analysis is performed by isolating DNA from an individual, manipulating the isolated DNA, *e.g.*, by digesting the DNA with restriction enzymes and/or amplifying a subset of sequences in the isolated DNA. The manipulated DNA is then examined further to determine if a particular sequence is present.

Commonly used procedures for analyzing the DNA include electrophoresis. Common applications of electrophoresis include agarose or polyacrylamide gel electrophoresis. DNA sequences are inserted, or loaded, on the gels and subjected to an electric field. Because DNA carries a uniform negative charge, DNA will migrate through the gel based on properties including sequence length, three-dimensional conformation and interactions with the gel matrix upon application of the electrical field. In most applications, smaller DNA molecules will migrate more rapidly through the gel than larger fragments. After electrophoresis has been continued for a sufficient length of time, the DNA molecules in the initial population of DNA sequences will have been separated according to their relative sizes.

Particular DNA molecules can then be detected using a variety of detection methodologies. For some applications, particular DNA sequences are identified by the presence of detectable tags, such as radioactive labels, attached to specific DNA molecules.

Electrophoretic-based separation analyses can be less desirable for applications in which it is desirable to rapidly, economically, and accurately analyze a large number of nucleic acid samples for particular sequence polymorphisms. For example, electrophoretic-based analysis can require a large amount of input DNA. In addition, processing the large number of samples required for electrophoretic-based nucleic acid based analyses can be labor intensive. Furthermore, these techniques can require samples of identical DNA molecules, which must be created prior to electrophoresis at costs that can be considerable.

Recently, automated electrophoresis systems have become available. However, electrophoresis can be ill-suited for applications such as clinical sequencing, where relatively cost-effective units with high throughput are needed. Thus, the need for non-electrophoretic methods for sequencing is great. For many applications, electrophoresis is used in conjunction with DNA sequence analysis.

Several alternatives to electrophoretic-based sequencing have been described. These include scanning tunnel electron microscopy, sequencing by hybridization, and single molecule detection methods.

Another alternative to electrophoretic-based separation analysis is solid substrate-based nucleic acid analyses. These methods typically rely upon the use of large numbers of nucleic acid probes affixed to different locations on a solid support. These solid supports can include, *e.g.*, glass surfaces, plastic microtiter plates, plastic sheets, thin polymers, or semi-conductors. The probes can be, *e.g.*, adsorbed or covalently attached to the support, or can be microencapsulated or otherwise entrapped within a substrate matrix, membrane, or film.

Substrate-based nucleic acid analyses can include applying a sample nucleic acid known or suspected of containing a particular sequence polymorphism to an array of probes attached to the solid substrate. The nucleic acids in the population are allowed to hybridize to complementary sequences attached to the substrate, if present. Hybridizing nucleic acid sequences are then detected in a detection step.

Solid support matrix-based hybridization and sequencing methodologies can require a high sample-DNA concentration and can be hampered by the relatively slow hybridization kinetics of nucleic acid samples with immobilized oligonucleotide probes. Often, only a small amount of template DNA is available, and it can be desirable to have high concentrations of the target nucleic acid sequence. Thus, substrate based detection analyses often include a step in which copies of the target nucleic acid, or a subset of sequences in the target nucleic acid, is amplified. Methods based on the Polymerase Chain Reaction (PCR), *e.g.*, can increase a small number of probe targets by several orders of magnitude in solution. However, PCR can be difficult to incorporate into a solid-phase approach because the amplified DNA is not immobilized onto the surface of the solid support matrix.

Solid-phase based detection of sequence polymorphisms has been described. An example is a "mini-sequencing" protocol based upon a solid phase principle described by Hultman, *et al.*, 1988. *Nucl. Acid. Res.* 17: 4937-4946; Syvanen, *et al.*, 1990. *Genomics* 8: 684-692. In this study, the incorporation of a radiolabeled nucleotide was measured and used for analysis of a three-allelic polymorphism of the human apolipoprotein E gene. However, such radioactive methods are not well-suited for routine clinical applications, and hence the development of a

simple, highly sensitive non-radioactive method for rapid DNA sequence analysis has also been of great interest.

SUMMARY OF THE INVENTION

The invention is based in part on the discovery of a highly sensitive method for determining the sequences of nucleic acids attached to solid substrates, and of novel substrate services for analyzing nucleic acid sequences.

Accordingly, in one aspect, the invention includes a substrate for analyzing a nucleic acid. The substrate includes a fiber optic surface onto which has been affixed one or more nucleic acid sequences. The fiber optic surface can be cavitated, *e.g.*, by etching the core of a fiber optic. The substrate can in addition include a plurality of bundled fiber optic surfaces, where one or more of the surfaces have anchored primers.

In another aspect, the invention includes an apparatus for analyzing a nucleic acid sequence. The apparatus can include a reagent delivery chamber, *e.g.*, a perfusion chamber, wherein the chamber includes a nucleic acid substrate, a conduit in communication with the perfusion chamber, an imaging system, *e.g.*, a fiber optic system, in communication with the perfusion chamber; and a data collection system in communication with the imaging system. The substrate can be a planar substrate. In other embodiments, the substrate can be the aforementioned fiber optic surface having nucleic acid sequences affixed to its termini.

In a further aspect, the invention includes a method for sequencing a nucleic acid. The method includes providing a primed anchor primer-circular template complex and combining the complex with a polymerase and nucleotides to generate concatenated, linear, complementary copies of the circular template (CLCC). The CLCC can be generated in solution and then linked to a solid substrate. Alternatively, one or more nucleic acid anchor primers can be linked to a solid support and then annealed to a plurality of circular nucleic acid templates which are then extended to yield a plurality of CLCCs.

One or more sequencing primers is then annealed to the CLCC to yield a primed sequencing primer-CLCC complex. Annealing of the sequencing primer can occur prior to, or after, attachment of the extended CLCC to the solid substrate. The sequencing primer(s) is then extended with a polymerase and a predetermined nucleotide triphosphate to yield a sequencing

product and a sequencing reaction byproduct, *e.g.*, inorganic pyrophosphate (PPi). If the predetermined nucleotide is incorporated into the primer, the sequencing reaction byproduct is generated and then identified, thereby determining the sequence of the nucleic acid. If the predetermined nucleotide is incorporated in the sequencing primer multiple times, *e.g.*, the concatenated nucleic acid template has multiple identical nucleotides, the quantity or concentration of sequencing reaction byproduct is measured to determine the number of nucleotides incorporated. If desired, additional predetermined nucleotide triphosphates can be added, *e.g.*, sequentially, and the presence or absence of sequence byproducts associated with each reaction can be determined.

In a still further aspect, the invention includes a method for sequencing a nucleic acid by providing one or more nucleic acid templates linked (*e.g.*, annealed) to a plurality of anchor primers linked to a fiber optic surface substrate, *e.g.*, the solid substrate discussed above.

In various embodiments of the apparatuses and methods described herein, the solid substrate includes two or more anchoring primers separated by approximately 10 μm to approximately 200 μm, 50 μm to approximately 150 μm, 100 μm to approximately 150 μm, or 150 μm. The solid support matrix can include a plurality of pads that are linked to the solid support. The surface area of the pads can be, *e.g.*, 10 μm^2 and one or more pads can be separated from one another by a distance ranging from approximately 50 μm to approximately 150 μm.

In preferred embodiments, at least a portion of the circular nucleic acid template is single-stranded DNA. The circular nucleic acid template can be, *e.g.*, genomic DNA or RNA, or a cDNA copy thereof. The circular nucleic acid can be, *e.g.*, 10-10,000, 20-1000, 10-200, 10-100, 10-50, or 20-40 nucleotides in length.

In some embodiments, multiple copies of one or more circular nucleic acids in the population are generated by a polymerase chain reaction. In other embodiments, the primed circular template is extended by rolling circle amplification (RCA). If desired, the template amplified by RCA can be further amplified by annealing a reverse primer to the single-stranded concatamer to yield a primed concatamer template and combining the primed concatamer template with a polymerase enzyme to generate multiple copies of the concatamer template. In still further embodiments, the template can be extended by a combination of PCR and RCA-amplification.

In preferred embodiments, the sequencing byproduct analyzed is pyrophosphate. When pyrophosphate is the detected byproduct, through its conversion to ATP, it is preferred that a dATP analog, *e.g.* α-thio dATP, be used by the polymerase in extending the annealed sequencing primer.

Preferably, the pyrophosphate is detected by contacting the sequencing byproduct with ATP sulfurylase under conditions sufficient to form ATP. The ATP can then be detected, *e.g.*, with an enzyme which generates a detectable product upon reaction with ATP. A preferred enzyme for detecting the ATP is luciferase. If desired, a wash buffer can be used between additions of various reactants herein. Preferably, apyrase is used to remove, *e.g.*, unreacted dNTP used to extend the sequencing primer. The wash buffer can optionally include apyrase.

The reactants and enzymes used herein, *e.g.*, the ATP sulfurylase, luciferase, and apyrase, can be attached to the solid surface.

The anchor primer sequence can include a linkage group, *e.g.* biotin, which can link the anchor primer to the solid support via a complementary group, *e.g.* avidin, attached to the solid support. In some embodiments, the solid support includes at least one optical fiber.

The invention also provides a method for profiling the concentrations of mRNA transcripts present in a cell. The identity of a transcript may be determined by the sequence at its 3' terminus (additional fragments may be used to distinguish between splice variants with identical 3' sequence). A sequencing apparatus having 10,000 (or more) sites could, in a single run, determine the mRNA species present at a concentration of as little as 1:10,000 (or less). Multiple runs, or multiple devices, could readily extend the limit to 1:100,000 or 1:1,000,000. This performance would be superior to current technologies, such as microarray hybridization, which have detection limits in the range 1:10,000 to 1:100,000.

In a further embodiment, the sequence of the amplified nucleic acid can be determined using byproducts of RNA synthesis. In this embodiment, an RNA transcript is generated from a promoter sequence present in the circular nucleic acid template library. Suitable promoter sites and their cognate RNA polymerases include RNA polymerases from *E. coli*, the RNA polymerase from the bacteriophage T3, the RNA polymerase from the bacteriophage T7, the RNA polymerase from the bacteriophage SP6, and the RNA polymerases from the viral families of bromoviruses, tobamoviruses, tombusvirus, lentiviruses, hepatitis C-like viruses, and

picornaviruses. To determine the sequence of an RNA transcript, a predetermined NTP, *i.e.*, an ATP, CTP, GTP, or UTP, is incubated with the template in the presence of the RNA polymerase. Incorporation of the test NTP into a nascent RNA strand can be determined by assaying for the presence of PPi using the enzymatic detection discussed herein.

The disclosures of one or more embodiments of the invention are set forth in the accompanying description below. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, the preferred methods and materials are now described. Other features, objects, and advantages of the invention will be apparent from the description and from the claims. In the specification and the appended claims, the singular forms include plural referents unless the context clearly dictates otherwise. Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Unless expressly stated otherwise, the techniques employed or contemplated herein are standard methodologies well known to one of ordinary skill in the art. The examples of embodiments are for illustration purposes only. All patents and publications cited in this specification are incorporated by reference.

BRIEF DESCRIPTION OF THE DRAWINGS

FIGS. 1A-D are schematic illustrations of rolling circle based amplification using an anchor primer.

FIG. 2 is a drawing of a sequencing apparatus according to the present invention.

FIG. 3 is a drawing of a perfusion chamber according to the present invention.

FIG. 4 is a drawing of a cavitated fiber optic terminus of the present invention.

FIG. 5 is a tracing of a sequence output of a concatemeric template generated using rolling circle amplification.

DETAILED DESCRIPTION OF THE INVENTION

The invention provides methods of preparing nucleic acid sequences for subsequent analysis, *e.g.*, sequencing, as well as methods and apparatuses for sequencing nucleic acids.

The methods described herein include a sample preparation process that results in a solid substrate array containing a plurality of anchor primers covalently linked to a nucleic acid containing one or more copies complementary to a target nucleic acid. Formation of the covalently linked anchor primer and one or more copies of the target nucleic acid preferably occurs by annealing the anchor primer to a complementary region of a circular nucleic acid, and then extending the annealed anchor primer with a polymerase to result in formation of a nucleic acid containing one or more copies of a sequence complementary to the circular nucleic acid.

Attachment of the anchor primer to the solid substrate can occur before, during, or subsequent to extension of the annealed anchor primer. Thus, in one embodiment, one or more anchor primers are linked to the solid substrate, after which the anchor primer is annealed to a target nucleic acid and extended in the presence of a polymerase. Alternatively, in a second embodiment, an anchor primer is first annealed to a target nucleic acid, and a 3'OH terminus of the annealed anchor primer is extended with a polymerase. The extended anchor primer is then linked to the solid substrate. By varying the sequence of anchor primers, it is possible to specifically amplify distinct target nucleic acids present in a population of nucleic acids.

Sequences in the target nucleic acid can be identified in a number of ways. Preferably, a sequencing primer is annealed to the amplified nucleic acid and used to generate a sequencing product. The nucleotide sequence of the sequence product is then determined, thereby allowing for the determination of the nucleic acid.

The methods and apparatuses described herein allow for the determination of nucleic acid sequence information without the need for first cloning a nucleic acid. In addition, the method is highly sensitive and can be used to determine the nucleotide sequence of a template nucleic acid which is present in only a few copies in a starting population of nucleic acids. Further, the method can be used to determine simultaneously the sequences of a large number of nucleic acids.

The methods and apparatuses described are generally useful for any application in which the identification of any particular nucleic acid sequence is desired. For example, the methods allow for identification of single nucleotide polymorphisms (SNPs), haplotypes involving multiple SNPs or other polymorphisms on a single chromosome, and transcript profiling. Other uses include sequencing of artificial DNA constructs to confirm or elicit their primary sequence,

or to identify specific mutant clones from random mutagenesis screens, as well as to obtain the sequence of cDNA from single cells, whole tissues or organisms from any developmental stage or environmental circumstance in order to determine the gene expression profile from that specimen. In addition, the methods allow for the sequencing of PCR products and/or cloned DNA fragments of any size isolated from any source.

The methods of the present invention can be also used for the sequencing of DNA fragments generated by analytical techniques that probe higher order DNA structure by their differential sensitivity to enzymes, radiation or chemical treatment (*e.g.*, partial DNase treatment of chromatin), or for the determination of the methylation status of DNA by comparing sequence generated from a given tissue with or without prior treatment with chemicals that convert methyl-cytosine to thymine (or other nucleotide) as the effective base recognized by the polymerase. Further, the methods of the present invention can be used to assay cellular physiology changes occurring during development or senescence at the level of primary sequence.

Methods of Sequencing Nucleic Acids

Structure of Anchor Primers

Anchor primers in general include a stalk region and at least two contiguous adapter regions. The stalk region is present at the 5' end of the anchor primer and includes a region of nucleotides for attaching the anchor primer to the solid substrate.

The anchor primer in general includes a region which hybridizes to a complementary sequence present in one or more members of a population of nucleic acid sequences. In some embodiments, the anchor primer includes two adjoining regions which hybridize to complementary regions ligated to separate ends of a target nucleic acid sequence. This embodiment is illustrated in FIG.1, which is discussed in more detail below.

In some embodiments, the adapter regions in the anchor primers are complementary to non-contiguous regions of sequence present in a second nucleic acid sequence. Each adapter region, for example, can be homologous to each terminus of a fragment produced by digestion with one or more restriction endonucleases. The fragment can include, *e.g.*, a sequence known or suspected to contain a sequence polymorphism.

In another example, the anchor primer may contain two adapter regions that are homologous to a gapped, *i.e.*, non-contiguous because of a deletion of one or more nucleotides, region of a target nucleic acid sequence. When adapter regions having these sequences are used, an aligning oligonucleotide corresponding to the gapped sequence may be annealed to the anchor primer along with a population of template nucleic acid molecules.

The anchor primer may optionally contain additional elements, *e.g.*, one or more restriction enzyme recognition sites, RNA polymerase binding sites (*e.g.*, a T7 promoter site).

One or more of the adapter regions may include, *e.g.*, a restriction enzyme recognition site or sequences present in identified DNA sequences, *e.g.*, sequences present in known genes. One or more adapter regions may also include sequences known to flank sequence polymorphisms. Sequence polymorphisms include nucleotide substitutions, insertions, deletions, or other rearrangements which result in a sequence difference between two otherwise identical nucleic acid sequences. An example of a sequence polymorphism is a single nucleotide polymorphism (SNP).

Linking of Anchor Primers to a Solid Support

In general, any nucleic acid capable of base-pairing can be used as an anchor primer. In some embodiments, the anchor primer is an oligonucleotide. As utilized herein the term oligonucleotide includes linear oligomers of natural or modified monomers or linkages, *e.g.*, deoxyribonucleosides, ribonucleosides, anomeric forms thereof, peptide nucleic acids (PNAs), and the like, that are capable of specifically binding to a target polynucleotide by way of a regular pattern of monomer-to-monomer interactions. These types of interactions can include, *e.g.*, Watson-Crick type of base-pairing, base stacking, Hoogsteen or reverse-Hoogsteen types of base-pairing, or the like. Generally, the monomers are linked by phosphodiester bonds, or analogs thereof, to form oligonucleotides ranging in size from, *e.g.*, 3-200, 8-150, 10-100, 20-80, or 25-50 monomeric units. Whenever an oligonucleotide is represented by a sequence of letters, it is understood that the nucleotides are oriented in the 5' $\longrightarrow$ 3' direction, from left-to-right, and that the letter "A" donates deoxyadenosine, the letter "T" denotes thymidine, the letter "C" denotes deoxycytosine, and the letter "G" denotes deoxyguanosine, unless otherwise noted herein. The oligonucleotides of the present invention can include non-natural nucleotide analogs.

However, where, for example, processing by enzymes is required, or the like, oligonucleotides comprising naturally-occurring nucleotides are generally required for maintenance of biological function.

Any material can be used as the solid support material, as long as the surface allows for stable attachment of the primers and detection of nucleic acid sequences. The solid support material can be planar or can be cavitated, e.g., in a cavitated terminus of a fiber optic or in a microwell etched, molded, or otherwise micromachined into the planar surface, *e.g.* using techniques commonly used in the construction of microelectromechanical systems. See *e.g.*, Rai-Choudhury, HANDBOOK OF MICROLITHOGRAPHY, MICROMACHINING, AND MICROFABRICATION, VOLUME I: MICROLITHOGRAPHY, Volume PM39, SPIE Press (1997); Madou, , CRC Press (1997), Aoki, *Biotech. Histochem.* 67: 98-9 (1992); Kane *et al.*, *Biomaterials*. 20: 2363-76 (1999); Deng *et al.*, *Anal. Chem.* 72:3176-80 (2000); Zhu *et al.*, *Nat. Genet.* 26:283-9 (2000). In some embodiments, the solid support is optically transparent, *e.g.*, glass.

The anchor primer can be linked to the solid support to reside on, or within, the solid support. In some embodiments, the plurality of anchor primers is linked to the solid support so they are spaced at regular intervals within an array. The periodicity between primers is preferably greater than the root-mean-square distance that products of the sequencing reactions diffuse prior to detection or the optical resolving power of the detection system, both of which are described in more detail below. The distance between primers on a solid substrate can be, *e.g.*,10-400 μm , 50-150 μm, 100-150 μm, or 150 μm.

An array of attachment sites on the optically transparent solid support can be constructed using lithographic techniques commonly used in the construction of electronic integrated circuits as described in, *e.g.*, techniques for attachment described in U.S. Patent Nos. 5,5143,854, 5,445,934, 5,744,305, and 5, 800,992; Chee *et al.*, *Science* 274: 610-614 (1996); Fodor *et al.*, *Nature* 364: 555-556 (1993); Fodor *et al.*, *Science* 251: 767-773 (1991); Gushin, *et al.*, *Anal. Biochem.* 250: 203-211 (1997); Kinosita *et al.*, *Cell* 93: 21-24 (1998); Kato-Yamada *et al.*, *J. Biol. Chem.* 273: 19375-19377 (1998); and Yasuda *et al.*, *Cell* 93: 1117-1124 (1998). Photolithography and electron beam lithography sensitize the solid support or substrate with a linking group that allows attachment of a modified biomolecule (*e.g.*, proteins or nucleic acids). See *e.g.*, Service, *Science* 283: 27-28 (1999); Rai-Choudhury, HANDBOOK OF

MICROLITHOGRAPHY, MICROMACHINING, AND MICROFABRICATION, VOLUME I: MICROLITHOGRAPHY, Volume PM39, SPIE Press (1997). Alternatively, an array of sensitized sites can be generated using thin-film technology as described in Zasadzinski *et al.*, *Science* 263: 1726-1733 (1994). The contents of all of these patents and publications are incorporated by reference in their entirety.

Anchor primers are linked to the solid substrate at the sensitized sites. A region of a solid substrate containing a linked primer is an anchor pad. Thus, by specifying the sensitized states on the solid support, it is possible to form an array or matrix of anchor pads. The anchor pads can be, *e.g.*, small diameter spots etched at evenly spaced intervals on the solid support. The anchor pads can be located at the bottoms of the cavitations or wells if the substrate has been cavitated, etched, or otherwise micromachined.

The anchor primer can be attached to the solid support via a covalent or non-covalent interaction. In general, any linkage recognized in the art can be used. Examples of such linkages common in the art include any suitable metal (*e.g.*, Co^{2+}, Ni^{2+})-hexahistidine complex, a biotin binding protein, *e.g.*, NEUTRAVIDIN™ modified avidin (Pierce Chemicals, Rockford, IL), streptavidin/biotin, avidin/biotin, glutathione S-transferase (GST)/glutathione, monoclonal antibody/antigen, and maltose binding protein/maltose, and pluronic coupling technologies. Samples containing the appropriate tag are incubated with the sensitized substrate so that zero, one, or multiple molecules attach at each sensitized site.

One biotin-(strept-)avidin-based anchoring method uses a thin layer of a photoactivatable biotin analog dried onto a solid surface. (Hengsakul and Cass, 1996. *Bioconjugate Chem.* 7: 249-254). The biotin analog is then exposed to white light through a mask, so as to create defined areas of activated biotin. Avidin (or streptavidin) is then added and allowed to bind to the activated biotin. The avidin possesses free biotin binding sites which can be utilized to "anchor" the biotinylated oligonucleotides through a biotin-(strept-)avidin linkage.

Alternatively, the anchor primer can be attached to the solid support with a biotin derivative possessing a photo-removable protecting group. This moiety is covalently bound to bovine serum albumin (BSA), which is attached to the solid support, *e.g.*, a glass surface. See Pirrung and Huang, 1996. *Bioconjugate Chem.* 7: 317-321. A mask is then used to create activated biotin within the defined irradiated areas. Avidin may then be localized to the

irradiated area, with biotinylated DNA subsequently attached through a BSA-biotin-avidin-biotin link. If desired, an intermediate layer of silane is deposited in a self-assembled monolayer on a silicon dioxide silane surface that can be patterned to localize BSA binding in defined regions. See *e.g.*, Mooney, *et al.*, 1996. *Proc. Natl. Acad. Sci. USA* 93: 12287-12291.

In pluronic based attachment, the anchor primers are first attached to the termini of a polyethylene oxide-polypropylene oxide-polyethylene oxide triblock copolymer , which is also known as a pluronic compound. The pluronic moiety can be used to attach the anchor primers to a solid substrate.

Pluronics attach to hydrophobic surfaces by virtue of the reaction between the hydrophobic surface and the polypropylene oxide. The remaining polyethylene oxide groups extend off the surface, thereby creating a hydrophilic environment. Nitrilotriacetic acid (NTA) can be conjugated to the terminal ends of the polyethylene oxide chains to allow for hexahistidine tagged anchor primers to be attached. In another embodiment, pyridyl disulfide (PDS) can be conjugated to the ends of the polyethylene chains allowing for attachment of a thiolated anchor primer via a disulfide bond. In one preferred embodiment, Pluronic F108 (BASF Corp.) is used for the attachment.

Each sensitized site on a solid support is potentially capable of attaching multiple anchor primers. Thus, each anchor pad may include one or more anchor primers. It is preferable to maximize the number of pads that have only a single productive reaction center (*e.g.*, the number of pads that, after the extension reaction, have only a single sequence extended from the anchor primer). This can be accomplished by techniques which include, but are not limited to: (*i*) varying the dilution of biotinylated anchor primers that are washed over the surface; (*ii*) varying the incubation time that the biotinylated primers are in contact with the avidin surface; (*iii*) varying the concentration of open- or closed-circular template so that, on average, only one primer on each pad is extended to generate the sequencing template; or (*iv*) reducing the size of the anchor pad to approach single-molecule dimensions (< 1 μm) such that binding of one anchor inhibits or blocks the binding of another anchor (*e.g.* by photoactivation of a small spot); or (*v*) reducing the size of the anchor pad such that binding of one circular template inhibits or blocks the binding of a second circular template.

In some embodiments, each individual pad contains just one linked anchor primer. Pads having only one anchor primer can be made by performing limiting dilutions of a selected anchor primer on to the solid support such that, on average, only one anchor primer is deposited on each pad. The concentration of anchor primer to be applied to a pad can be calculated utilizing, for example, a Poisson distribution model.

In order to maximize the number of reaction pads that contain a single anchor primer, a series of dilution experiments are performed in which a range of anchor primer concentrations or circular template concentrations are varied. For highly dilute concentrations of primers, primers and circular templates binding to the same pad will be independent of each other, and a Poisson distribution will characterize the number of anchor primers extended on any one pad. Although there will be variability in the number of primers that are actually extended, a maximum of 37% of the pads will have a single extended anchor primer (the number of pads with a single anchor oligonucleotide). This number can be obtained as follows.

Let N_p be the average number of anchor primers on a pad and f be the probability that an anchor primer is extended with a circular template. Then the average number of extended anchor primers per pad is N_pf, which is defined as the quantity **a**. There will be variability in the number of primers that are actually extended. In the low-concentration limit, primers and circular templates binding to the same pad will be independent of each other, and a Poisson distribution P(n) will characterize the number of anchor primers n extended on any pad. This distribution may be mathematically defined by: $P(n) = (\mathbf{a}^n / n!)\exp(-\mathbf{a})$, with $P(1) = \mathbf{a}\exp(-\mathbf{a})$. The probability P(1) assumes its maximum value exp(-1) for a = 1, with 37% of pads having a single extended anchor primer.

A range of anchor primer concentrations and circular template concentrations may be subsequently scanned to find a value of N_pf closest to 1. A preferable method to optimize this distribution is to allow multiple anchor primers on each reaction pad, but use a limiting dilution of circular template so that, on average, only one primer on each pad is extended to generate the sequencing template.

Alternatively, at low concentrations of anchor primers, at most one anchor primer will likely be bound on each reaction pad. A high concentration of circular template may be used so that each primer is likely to be extended.

Where the reaction pads are arrayed on a planar surface or a fiber optic array, the individual pads are approximately 10 μm on a side, with a 100 μm spacing between adjacent pads. Hence, on a 1 cm^2 surface a total of approximately 10,000 microreactors could be deposited, and, according to the Poisson distribution, approximately 3700 of these will contain a single anchor primer. In certain embodiments, after the primer oligonucleotide has been attached to the solid support, modified, *e.g.*, biotinylated, enzymes are deposited to bind to the remaining, unused avidin binding sites on the surface.

In other embodiments multiple anchor primers are attached to any one individual pad in an array. Limiting dilutions of a plurality of circular nucleic acid templates (described in more detail below) may be hybridized to the anchor primers so immobilized such that, on average, only one primer on each pad is hybridized to a nucleic acid template. Library concentrations to be used may be calculated utilizing, for example, limiting dilutions and a Poisson distribution model.

Libraries of single-stranded circular templates

A plurality of nucleic acid templates, *e.g.*, a nucleic acid library, in general includes open circular or closed circular nucleic acid molecules. A "closed circle" is a covalently closed circular nucleic acid molecule, *e.g.*, a circular DNA or RNA molecule. An "open circle" is a linear single-stranded nucleic acid molecule having a 5' phosphate group and a 3' hydroxyl group. In some embodiments, the open circle is formed *in situ* from a linear double-stranded nucleic acid molecule. The ends of a given open circle nucleic acid molecule can be ligated by DNA ligase. Sequences at the 5' and 3' ends of the open circle molecule are complementary to two regions of adjacent nucleotides in a second nucleic acid molecule, *e.g.*, an adapter region of an anchor primer, or to two regions that are nearly adjoining in a second DNA molecule. Thus, the ends of the open-circle molecule can be ligated using DNA ligase, or extended by DNA polymerase in a gap-filling reaction. Open circles are described in detail in Lizardi, U.S. Pat. No. 5,854,033. An open circle can be converted to a closed circle in the presence of a DNA ligase (for DNA) or RNA ligase following, *e.g.*, annealing of the open circle to an anchor primer.

If desired, nucleic acid templates can be provided as padlock probes. Padlock probes are linear oligonucleotides that include target-complementary sequences located at each end, and

which are separated by a linker sequence. The linkers can be ligated to ends of members of a library of nucleic acid sequences that have been, *e.g.*, physically sheared or digested with restriction endonucleases. Upon hybridization to a target-sequence, the two ends of the probes are brought in juxtaposition, and they can then be joined through enzymatic ligation. The linkers can be ligated to ends of members of a library of nucleic acid sequences that have been, *e.g.*, physically sheared or digested with restriction endonucleases.

The 5'- and 3'-terminal regions of these linear oligonucleotides are designed to basepair adjacent to one another on a specific target sequence strand, thus the termini of the linear oligonucleotide are brought into juxtaposition by hybridization to the target sequence. This juxtaposition allows the two probe segments (if properly hybridized) to be covalently-bound by enzymatic ligation (*e.g.*, with T4 DNA ligase), thus converting the probes to circularly-closed molecules which are catenated to the specific target sequences (see *e.g.*, Nilsson, *et al.*, 1994. *Science* **265**: 2085-2088). The resulting probes are suitable for the simultaneous analysis of many gene sequences both due to their specificity and selectivity for gene sequence variants (see *e.g.*, Lizardi, *et al.*, 1998. Nat. Genet. **19**: 225-232; Nilsson, *et al.*, 1997. *Nat. Genet.* **16**: 252-255) and due to the fact that the resulting reaction products remain localized to the specific target sequences. Moreover, intramolecular ligation of many different probes is expected to be less susceptible to non-specific cross-reactivity than multiplex PCR-based methodologies where non-cognate pairs of primers can give rise to irrelevant amplification products (see *e.g.*, Landegren and Nilsson, 1997. *Ann. Med.* **29**: 585-590).

The starting library can be either single-stranded or double-stranded, as long as it includes a region that, if present in the library, is available for annealing, or can be made available for annealing, to an anchor primer sequence. When used as a template for rolling circle amplification, a region of the double-stranded template needs to be at least transiently single-stranded in order to act as a template for extension of the anchor primer.

Library templates can include multiple elements, including, but not limited to, one or more regions that are complementary to the anchor primer. For example, the template libraries may include a region complementary to a sequencing primer, a control nucleotide region, and an insert sequence comprised of the sequencing template to be subsequently characterized. As is explained in more detail below, the control nucleotide region is used to calibrate the relationship

between the amount of byproduct and the number of nucleotides incorporated. As utilized herein the term "complement" refers to nucleotide sequences that are able to hybridize to a specific nucleotide sequence to form a matched duplex.

In one embodiment, a library template includes: (*i*) two distinct regions that are complementary to the anchor primer, (*ii*) one region homologous to the sequencing primer, (*iii*) one optional control nucleotide region, (*iv*) an insert sequence of, *e.g.*, 30-500, 50-200, or 60-100 nucleotides, that is to be sequenced. The template can, of course, include two, three, or all four of these features.

The template nucleic acid can be constructed from any source of nucleic acid, *e.g.*, any cell, tissue, or organism, and can be generated by any art-recognized method. Suitable methods include, *e.g.*, sonication of genomic DNA and digestion with one or more restriction endonucleases (RE) to generate fragments of a desired range of lengths from an initial population of nucleic acid molecules. Preferably, one or more of the restriction enzymes have distinct four-base recognition sequences. Examples of such enzymes include, *e.g.*, Sau3A1, MspI, and TaqI. Preferably, the enzymes are used in conjunction with anchor primers having regions containing recognition sequences for the corresponding restriction enzymes. In some embodiments, one or both of the adapter regions of the anchor primers contain additional sequences adjoining known restriction enzyme recognition sequences, thereby allowing for capture or annealing to the anchor primer of specific restriction fragments of interest to the anchor primer.

In other embodiments, the restriction enzyme is used with a type IIS restriction enzyme.

Alternatively, template libraries can be made by generating a complementary DNA (cDNA) library from RNA, *e.g.*, messenger RNA (mRNA). The cDNA library can, if desired, be further processed with restriction endonucleases to obtain a 3' end characteristic of a specific RNA, internal fragments, or fragments including the 3' end of the isolated RNA. Adapter regions in the anchor primer may be complementary to a sequence of interest that is thought to occur in the template library, *e.g.*, a known or suspected sequence polymorphism within a fragment generated by endonuclease digestion.

In one embodiment, an indexing oligonucleotide can be attached to members of a template library to allow for subsequent correlation of a template nucleic acid with a population of nucleic acids from which the template nucleic acid is derived. For example, one or more

samples of a starting DNA population can be fragmented separately using any of the previously disclosed methods (*e.g.*, restriction digestion, sonication). An indexing oligonucleotide sequence specific for each sample is attached to, *e.g.*, ligated to, the termini of members of the fragmented population. The indexing oligonucleotide can act as a region for circularization, amplification and, optionally, sequencing, which permits it to be used to index, or code, a nucleic acid so as to identify the starting sample from which it is derived.

Distinct template libraries made with a plurality of distinguishable indexing primers can be mixed together for subsequent reactions. Determining the sequence of the member of the library allows for the identification of a sequence corresponding to the indexing oligonucleotide. Based on this information, the origin of any given fragment can be inferred.

Annealing and Amplification of Primer-Template Nucleic Acid Complexes

Libraries of nucleic acids are annealed to anchor primer sequences using recognized techniques (see, *e.g.*, Hatch, *et al.*, 1999. *Genet. Anal. Biomol. Engineer.* 15: 35-40; Kool, U.S. Patent No. 5,714, 320 and Lizardi, U.S. Patent No. 5,854,033). In general, any procedure for annealing the anchor primers to the template nucleic acid sequences is suitable as long as it results in formation of specific, *i.e.*, perfect or nearly perfect, complementarity between the adapter region or regions in the anchor primer sequence and a sequence present in the template library.

A number of *in vitro* nucleic acid amplification techniques may be utilized to extend the anchor primer sequence. The size of the amplified DNA preferably is smaller than the size of the anchor pad and also smaller than the distance between anchor pads.

The amplification is typically performed in the presence of a polymerase, *e.g.*, a DNA or RNA-directed DNA polymerase, and one, two, three, or four types of nucleotide triphosphates, and, optionally, auxiliary binding proteins. In general, any polymerase capable of extending a primed 3'-OH group can be used a long as it lacks a 3' to 5' exonuclease activity. Suitable polymerases include, *e.g.*, the DNA polymerases from *Bacillus stearothermophilus*, *Thermus acquaticus*, *Pyrococcus furiosis*, *Thermococcus litoralis*, and *Thermus thermophilus*, bacteriophage T4 and T7, and the *E. coli* DNA polymerase I Klenow fragment. Suitable RNA-directed DNA polymerases include, *e.g.*, the reverse transcriptase from the Avian Myeloblastosis

Virus, the reverse transcriptase from the Moloney Murine Leukemia Virus, and the reverse transcriptase from the Human Immunodeficiency Virus-I.

A number of *in vitro* nucleic acid amplification techniques have been described. These amplification methodologies may be differentiated into those methods: (*i*) which require temperature cycling - polymerase chain reaction (PCR) (see *e.g.*, Saiki, *et al.*, 1995. *Science* **230**: 1350-1354), ligase chain reaction (see *e.g.*, Barany, 1991. *Proc. Natl. Acad. Sci. USA* **88**: 189-193; Barringer, *et al.*, 1990. *Gene* **89**: 117-122) and transcription-based amplification (see *e.g.*, Kwoh, *et al.*, 1989. *Proc. Natl. Acad. Sci. USA* **86**: 1173-1177) and (*ii*) isothermal amplification systems - self-sustaining, sequence replication (see *e.g.*, Guatelli, *et al.*, 1990. *Proc. Natl. Acad. Sci. USA* **87**: 1874-1878); the Qβ replicase system (see *e.g.*, Lizardi, *et al.*, 1988. *BioTechnology* **6**: 1197-1202); strand displacement amplification Nucleic Acids Res. 1992 Apr 11;20(7):1691-6.; and the methods described in PNAS 1992 Jan 1;89(1):392-6; and NASBA J Virol Methods. 1991 Dec;35(3):273-86.

Isothermal amplification also includes rolling circle-based amplification (RCA). RCA is discussed in, *e.g.*, Kool, U.S. Patent No. 5,714,320 and Lizardi, U.S. Patent No. 5,854,033; Hatch, *et al.*, 1999. *Genet. Anal. Biomol. Engineer.* 15: 35-40. The result of the RCA is a single DNA strand extended from the 3' terminus of the anchor primer (and thus is linked to the solid support matrix) and including a concatamer containing multiple copies of the circular template annealed to a primer sequence. Typically, 1,000 to 10,000 or more copies of circular templates, each having a size of, *e.g.*, approximately 30-500, 50-200, or 60-100 nucleotides size range, can be obtained with RCA.

The product of RCA amplification following annealing of a circular nucleic acid molecule to an anchor primer is shown schematically in FIG. 1A. A circular template nucleic acid 102 is annealed to an anchor primer 104, which has been linked to a surface 106 at its 5' end and has a free 3' OH available for extension. The circular template nucleic acid 102 includes two adapter regions 108 and 110 which are complementary to regions of sequence in the anchor primer 104. Also included in the circular template nucleic acid 102 is an insert 112 and a region 114 homologous to a sequencing primer, which is used in the sequencing reactions described below.

Upon annealing, the free 3'-OH on the anchor primer 104 can be extended using sequences within the template nucleic acid 102. The anchor primer 102 can be extended along the template multiple times, with each iteration adding to the sequence extended from the anchor primer a sequence complementary to the circular template nucleic acid. Four iterations, or four rounds of rolling circle replication, are shown in FIG.1A as the extended anchor primer amplification product 114. Extension of the anchor primer results in an amplification product covalently or otherwise physically attached to the substrate 106.

Additional embodiments of circular templates and anchor primers are shown in more detail in FIGS. 1B-1F. FIG. 1B illustrates an annealed open circle linear substrate that can serve, upon ligation, as a template for extension of an anchor primer. A template molecule having the sequence 5' – TCg TgT gAg gTC TCA gCA TCT TAT gTA TAT TTA CTT CTA TTC TCA gTT gCC TAA gCT gCA gCC A – 3' (SEQ ID NO:1) is annealed to an anchor primer having a biotin linker at its 5' terminus and the sequence 5'-gAC CTC ACA CgA Tgg CTg CAg CTT – 3' (SEQ ID NO:2). Annealing of the template results in juxtaposition of the 5' and 3' ends of the template molecule. The 3'OH of the anchor primer can be extended using the circular template.

The use of a circular template and an anchor primer for identification of single nucleotide polymorphisms is shown in FIG. 1C. Shown is a generic anchor primer having the sequence 5' - gAC CTC ACA CgA Tgg CTg CAg CTT – 3'(SEQ ID NO:3). The anchor primer anneals to an SNP probe having the sequence 5' – TTT ATA TgT ATT CTA CgA CTC Tgg AgT gTg CTA CCg ACg TCg AAt CCg TTg ACT CTT ATC TTC A – 3' (SEQ ID NO:4). The SNP probe in turn hybridizes to a region of a SNP-containing region of a gene having the sequence 5' – CTA gCT CgT ACA TAT AAA TgA AgA TAA gAT CCT g – 3' (SEQ ID NO:5). Hybridization of a nucleic acid sequence containing the polymorphism to the SNP probe complex allows for subsequent ligation and circularization of the SNP probe. The SNP probe is designed so that its 5' and 3' termini anneal to the genomic region so as to abut in the region of the polymorphic site, as is indicated in FIG. 1C. The circularized SNP probe can be subsequently extended and sequenced using the methods described herein. A nucleic acid lacking the polymorphism does not hybridize so as to result in juxtaposition of the 5' and 3'

termini of the SNP probe. In this case, the SNP probe cannot be ligated to form a circular substrate needed for subsequent extension.

FIG. 1D illustrates the use of a gap oligonucleotide to along with a circular template molecule. An anchor primer having the sequence 5'–gAC CTC ACA CgA gTA gCA Tgg CTg CAg CTT – 3' (SEQ ID NO:6) is attached to a surface through a biotin linker. A template molecule having the sequence 5' – TCg TgT gAg gTC TCA gCA TCT TAT gTA TAT TTA CTT CTA TTC TCA gTT gCC TAA gCT gCA gCC A – 3' (SEQ ID NO:7) is annealed to the anchor primer to result in partially single stranded, or gapped region, in the anchor primer flanked by a double-stranded region. A gapping molecule having the sequence 5' – TgC TAC – 3' then anneals to the anchor primer. Ligation of both ends of the gap oligonucleotide to the template molecule results in formation of a circular nucleic acid molecule that can act as a template for rolling circle amplification.

Circular oligonucleotides that are generated during polymerase-mediated DNA replication are dependent upon the relationship between the template and the site of replication initiation. In double-stranded DNA templates, the critical features include whether the template is linear or circular in nature, and whether the site of initiation of replication (*i.e.*, the replication "fork") is engaged in synthesizing both strands of DNA or only one. In conventional double-stranded DNA replication, the replication fork is treated as the site at which the new strands of DNA are synthesized. However, in linear molecules (whether replicated unidirectionally or bidirectionally), the movement of the replication fork(s) generate a specific type of structural motif. If the template is circular, one possible spatial orientation of the replicating molecule takes the form of a θ structure.

Alternatively, RCA can occur when the replication of the duplex molecule begins at the origin. Subsequently, a nick opens one of the strands, and the free 3'-terminal hydroxyl moiety generated by the nick is extended by the action of DNA polymerase. The newly synthesized strand eventually displaces the original parental DNA strand. This aforementioned type of replication is known as rolling-circle replication (RCR) because the point of replication may be envisaged as "rolling around" the circular template strand and, theoretically, it could continue to do so indefinitely. Additionally, because the newly synthesized DNA strand is covalently-bound to the original template, the displaced strand possesses the original genomic sequence (*e.g.*, gene

or other sequence of interest) at its 5'-terminus. In rolling-circle replication, the original genomic sequence is followed by any number of "replication units" complementary to the original template sequence, wherein each replication unit is synthesized by continuing revolutions of said original template sequence. Hence, each subsequent revolution displaces the DNA which is synthesized in the previous replication cycle.

In vivo, rolling-circle replication is utilized in several biological systems. For example, the genome of several bacteriophage are single-stranded, circular DNA. During replication, the circular DNA is initially converted to a duplex form, which is then replicated by the aforementioned rolling-circle replication mechanism. The displaced terminus generates a series of genomic units that can be cleaved and inserted into the phage particles. Additionally, the displaced single-strand of a rolling-circle can be converted to duplex DNA by synthesis of a complementary DNA strand. This synthesis can be used to generate the concatemeric duplex molecules required for the maturation of certain phage DNAs. For example, this provides the principle pathway by which λ bacteriophage matures. Rolling-circle replication is also used *in vivo* to generate amplified rDNA in *Xenopus* oocytes, and this fact may help explain why the amplified rDNA is comprised of a large number of identical repeating units. In this case, a single genomic repeating unit is converted into a rolling-circle. The displaced terminus is then converted into duplex DNA which is subsequently cleaved from the circle so that the two termini can be ligated together so as to generate the amplified circle of rDNA.

Through the use of the RCA reaction, a strand may be generated which represents many tandem copies of the complement to the circularized molecule. For example, RCA has recently been utilized to obtain an isothermal cascade amplification reaction of circularized padlock probes *in vitro* in order to detect single-copy genes in human genomic DNA samples (see Lizardi, *et al.*, 1998. *Nat. Genet.* **19**: 225-232). In addition, RCA has also been utilized to detect single DNA molecules in a solid phase-based assay, although difficulties arose when this technique was applied to *in situ* hybridization (see Lizardi, *et al.*, 1998. *Nat. Genet.* **19**: 225-232).

If desired, RCA can be performed at elevated temperatures, *e.g.*, at temperatures greater than 37° C, 42° C, 45° C, 50° C, 60° C, or 70° C. In addition, RCA can be performed initially at a lower temperature, *e.g.*, room temperature, and then shifted to an elevated temperature.

Elevated temperature RCA is preferably performed with thermostable nucleic acid polymerases and with primers that can anneal stably and with specificity at elevated temperatures.

RCA can also be performed with non-naturally occurring oligonucleotides, *e.g.*, peptide nucleic acids. Further, RCA can be performed in the presence of auxiliary proteins such as single-stranded binding proteins.

The development of a method of amplifying short DNA molecules which have been immobilized to a solid support, termed rolling circle amplification (RCA) has been recently described in the literature (see *e.g.*, Hatch, *et al.*, 1999. Rolling circle amplification of DNA immobilized on solid surfaces and its application to multiplex mutation detection. *Genet. Anal. Biomol. Engineer.* **15**: 35-40; Zhang, *et al.*, 1998. Amplification of target-specific, ligation-dependent circular probe. *Gene* **211**: 277-85; Baner, *et al.*, 1998. Signal amplification of padlock probes by rolling circle replication. *Nucl. Acids Res.* **26**: 5073-5078; Liu, *et al.*, 1995. Rolling circle DNA synthesis: small circular oligonucleotides as efficient templates for DNA polymerase. *J. Am. Chem. Soc.* **118**: 1587-1594; Fire and Xu, 1995. Rolling replication of short DNA circles. *Proc. Natl. Acad. Sci. USA* **92**: 4641-4645; Nilsson, *et al.*, 1994. Padlock probes: circularizing oligonucleotides for localized DNA detection. *Science* **265**: 2085-2088). RCA targets specific DNA sequences through hybridization and a DNA ligase reaction. The circular product is then subsequently used as a template in a rolling circle replication reaction.

Rolling-circle amplification (RCA) driven by DNA polymerase can replicate circularized oligonucleotide probes with either linear or geometric kinetics under isothermal conditions. In the presence of two primers (one hybridizing to the + strand, and the other, to the - strand of DNA), a complex pattern of DNA strand displacement ensues which possesses the ability to generate $1x10^9$ or more copies of each circle in a short period of time (*i.e.*, less-than 90 minutes), enabling the detection of single-point mutations within the human genome. Using a single primer, RCA generates hundreds of randomly-linked copies of a covalently closed circle in several minutes. If solid support matrix-associated, the DNA product remains bound at the site of synthesis, where it may be labeled, condensed, and imaged as a point light source. For example, linear oligonucleotide probes, which can generate RCA signals, have been bound covalently onto a glass surface. The color of the signal generated by these probes indicates the

allele status of the target, depending upon the outcome of specific, target-directed ligation events. As RCA permits millions of individual probe molecules to be counted and sorted, it is particularly amenable for the analysis of rare somatic mutations. RCA also shows promise for the detection of padlock probes bound to single-copy genes in cytological preparations.

In addition, a solid-phase RCA methodology has also been developed to provide an effective method of detecting constituents within a solution. Initially, a recognition step is used to generate a complex h a circular template is bound to a surface. A polymerase enzyme is then used to amplify the bound complex. RCA uses small DNA probes that are amplified to provide an intense signal using detection methods, including the methods described in more detail below.

Other examples of isothermal amplification systems include, *e.g.*, (*i*) self-sustaining, sequence replication (see *e.g.*, Guatelli, *et al.*, 1990. *Proc. Natl. Acad. Sci. USA* 87: 1874-1878), (*ii*) the Qβ replicase system (see *e.g.*, Lizardi, *et al.*, 1988. *BioTechnology* 6: 1197-1202), and (*iii*) nucleic acid sequence-based amplification (NASBA™; see Kievits, *et al.*, 1991. *J. Virol. Methods* 35: 273-286).

Determining the nucleotide sequence of the sequence product

Amplification of a nucleic acid template as described above results in multiple copies of a template nucleic acid sequence covalently linked to an anchor primer. In one embodiment, a region of the sequence product is determined by annealing a sequencing primer to a region of the template nucleic acid, and then contacting the sequencing primer with a DNA polymerase and a known nucleotide triphosphate, *i.e.*, dATP, dCTP, dGTP, dTTP, or an analog of one of these nucleotides. The sequence can be determined by detecting a sequence reaction byproduct, as is described below.

The sequence primer can be any length or base composition, as long as it is capable of specifically annealing to a region of the amplified nucleic acid template. No particular structure for the sequencing primer is required so long as it is able to specifically prime a region on the amplified template nucleic acid. Preferably, the sequencing primer is complementary to a region of the template that is between the sequence to be characterized and the sequence hybridizable to the anchor primer. The sequencing primer is extended with the DNA polymerase to form a

sequence product. The extension is performed in the presence of one or more types of nucleotide triphosphates, and if desired, auxiliary binding proteins.

Incorporation of the dNTP is preferably determined by assaying for the presence of a sequencing byproduct. In a preferred embodiment, the nucleotide sequence of the sequencing product is determined by measuring inorganic pyrophosphate (PPi) liberated from a nucleotide triphosphate (dNTP) as the dNMP is incorporated into an extended sequence primer. This method of sequencing, termed Pyrosequencing™ technology (PyroSequencing AB, Stockholm, Sweden) can be performed in solution (liquid phase) or as a solid phase technique. PPi-based sequencing methods are described generally in, *e.g.,* WO9813523A1, Ronaghi, *et al.,* 1996. *Anal. Biochem.* 242: 84-89, and Ronaghi, *et al.*, 1998. *Science* 281: 363-365 (1998). These disclosures of PPi sequencing are incorporated herein in their entirety, by reference.

Pyrophosphate released under these conditions can be detected enzymatically (*e.g.,* by the generation of light in the luciferase-luciferin reaction). Such methods enable a nucleotide to be identified in a given target position, and the DNA to be sequenced simply and rapidly while avoiding the need for electrophoresis and the use of potentially dangerous radiolabels.

PPi can be detected by a number of different methodologies, and various enzymatic methods have been previously described (see *e.g.,* Reeves, *et al.,* 1969. *Anal. Biochem.* **28**: 282-287; Guillory, *et al.,* 1971. *Anal. Biochem.* **39**: 170-180; Johnson, *et al.,* 1968. *Anal. Biochem.* **15**: 273; Cook, *et al.,* 1978. *Anal. Biochem.* **91**: 557-565; and Drake, *et al.,* 1979. *Anal. Biochem.* **94**: 117-120).

PPi liberated as a result of incorporation of a dNTP by a polymerase can be converted to ATP using, *e.g.*, an ATP sulfurylase. This enzyme has been identified as being involved in sulfur metabolism. Sulfur, in both reduced and oxidized forms, is an essential mineral nutrient for plant and animal growth (see *e.g.,* Schmidt and Jager, 1992. *Ann. Rev. Plant Physiol. Plant Mol. Biol.* **43**: 325-349). In both plants and microorganisms, active uptake of sulfate is followed by reduction to sulfide. As sulfate has a very low oxidation/reduction potential relative to available cellular reductants, the primary step in assimilation requires its activation via an ATP-dependent reaction (see *e.g.,* Leyh, 1993. *Crit. Rev. Biochem. Mol. Biol.* **28**: 515-542). ATP sulfurylase (ATP: sulfate adenylyltransferase; EC 2.7.7.4) catalyzes the initial reaction in the metabolism of inorganic sulfate (SO_4^{-2}); see *e.g.,* Robbins and Lipmann, 1958. *J. Biol. Chem.*

233: 686-690; Hawes and Nicholas, 1973. *Biochem. J.* **133**: 541-550). In this reaction SO_4^{-2} is activated to adenosine 5'-phosphosulfate (APS).

ATP sulfurylase has been highly purified from several sources, such as *Saccharomyces cerevisiae* (see *e.g.*, Hawes and Nicholas, 1973. *Biochem. J.* **133**: 541-550); *Penicillium chrysogenum* (see *e.g.*, Renosto, *et al.*, 1990. *J. Biol. Chem.* **265**: 10300-10308); rat liver (see *e.g.*, Yu, *et al.*, 1989. *Arch. Biochem. Biophys.* **269**: 165-174); and plants (see *e.g.*, Shaw and Anderson, 1972. *Biochem. J.* **127**: 237-247; Osslund, *et al.*, 1982. *Plant Physiol.* **70**: 39-45). Furthermore, ATP sulfurylase genes have been cloned from prokaryotes (see *e.g.*, Leyh, *et al.*, 1992. *J. Biol. Chem.* **267**: 10405-10410; Schwedock and Long, 1989. *Mol. Plant Microbe Interaction* **2**: 181-194; Laue and Nelson, 1994. *J. Bacteriol.* **176**: 3723-3729); eukaryotes (see *e.g.*, Cherest, *et al.*, 1987. *Mol. Gen. Genet.* **210**: 307-313; Mountain and Korch, 1991. *Yeast* **7**: 873-880; Foster, *et al.*, 1994. *J. Biol. Chem.* **269**: 19777-19786); plants (see *e.g.*, Leustek, *et al.*, 1994. *Plant Physiol.* **105**: 897-90216); and animals (see *e.g.*, Li, *et al.*, 1995. *J. Biol. Chem.* **270**: 29453-29459). The enzyme is a homo-oligomer or heterodimer, depending upon the specific source (see *e.g.*, Leyh and Suo, 1992. *J. Biol. Chem.* **267**: 542-545).

In some embodiments, a thermostable sulfurylase is used. Thermostable sulfurylases can be obtained from, *e.g.*, Archaeoglobus or Pyrococcus spp. Sequences of thermostable sulfurylases are available at database Acc. No. 028606, Acc. No. Q9YCR4, and Acc. No. P56863.

ATP sulfurylase has been used for many different applications, for example, bioluminometric detection of ADP at high concentrations of ATP (see *e.g.*, Schultz, *et al.*, 1993. *Anal. Biochem.* **215**: 302-304); continuous monitoring of DNA polymerase activity (see *e.g.*, Nyrbn, 1987. *Anal. Biochem.* **167**: 235-238); and DNA sequencing (see *e.g.*, Ronaghi, *et al.*, 1996. *Anal. Biochem.* **242**: 84-89; Ronaghi, *et al.*, 1998. *Science* **281**: 363-365; Ronaghi, *et al.*, 1998. *Anal. Biochem.* **267**: 65-71).

Several assays have been developed for detection of the forward ATP sulfurylase reaction. The colorimetric molybdolysis assay is based on phosphate detection (see *e.g.*, Wilson and Bandurski, 1958. *J. Biol. Chem.* **233**: 975-981), whereas the continuous spectrophotometric molybdolysis assay is based upon the detection of NADH oxidation (see *e.g.*, Seubert, *et al.*,

1983. *Arch. Biochem. Biophys.* **225**: 679-691; Seubert, *et al.*, 1985. *Arch. Biochem. Biophys.* **240**: 509-523). The later assay requires the presence of several detection enzymes. In addition, several radioactive assays have also been described in the literature (see *e.g.*, Daley, *et al.*, 1986. *Anal. Biochem.* **157**: 385-395). For example, one assay is based upon the detection of ^{32}PPi released from ^{32}P-labeled ATP (see *e.g.*, Seubert, *et al.*, 1985. *Arch. Biochem. Biophys.* **240**: 509-523) and another on the incorporation of ^{35}S into [^{35}S]-labeled APS (this assay also requires purified APS kinase as a coupling enzyme; see *e.g.*, Seubert, *et al.*, 1983. *Arch. Biochem. Biophys.* **225**: 679-691); and a third reaction depends upon the release of $^{35}SO_4^{-2}$ from [^{35}S]-labeled APS (see *e.g.*, Daley, *et al.*, 1986. *Anal. Biochem.* **157**: 385-395).

For detection of the reversed ATP sulfurylase reaction a continuous spectrophotometric assay (see *e.g.*, Segel, *et al.*, 1987. *Methods Enzymol.* **143**: 334-349); a bioluminometric assay (see *e.g.*, Balharry and Nicholas, 1971. *Anal. Biochem.* **40**: 1-17); an $^{35}SO_4^{-2}$ release assay (see *e.g.*, Seubert, *et al.*, 1985. *Arch. Biochem. Biophys.* **240**: 509-523); and a ^{32}PPi incorporation assay (see *e.g.*, Osslund, *et al.*, 1982. *Plant Physiol.* **70**: 39-45) have been previously described.

ATP produced by an ATP sulfurylase can be hydrolyzed using enzymatic reactions to generate light. Light-emitting chemical reactions (*i.e.*, chemiluminescence) and biological reactions (*i.e.*, bioluminescence) are widely used in analytical biochemistry for sensitive measurements of various metabolites. In bioluminescent reactions, the chemical reaction that leads to the emission of light is enzyme-catalyzed. For example, the luciferin-luciferase system allows for specific assay of ATP and the bacterial luciferase-oxidoreductase system can be used for monitoring of NAD(P)H. Both systems have been extended to the analysis of numerous substances by means of coupled reactions involving the production or utilization of ATP or NAD(P)H (see *e.g.*, Kricka, 1991. Chemiluminescent and bioluminescent techniques. *Clin. Chem.* **37**: 1472-1281).

The development of new reagents have made it possible to obtain stable light emission proportional to the concentrations of ATP (see *e.g.*, Lundin, 1982. Applications of firefly luciferase In; *Luminescent Assays* (Raven Press, New York) or NAD(P)H (see *e.g.*, Lovgren, *et al.*, Continuous monitoring of NADH-converting reactions by bacterial luminescence. *J. Appl. Biochem.* **4**: 103-111). With such stable light emission reagents, it is possible to make endpoint

assays and to calibrate each individual assay by addition of a known amount of ATP or NAD(P)H. In addition, a stable light-emitting system also allows continuous monitoring of ATP- or NAD(P)H-converting systems.

Suitable enzymes for converting ATP into light include luciferases, *e.g.*, insect luciferases. Luciferases produce light as an end-product of catalysis. The best known light-emitting enzyme is that of the firefly, *Photinus pyralis* (*Coleoptera*). The corresponding gene has been cloned and expressed in bacteria (see *e.g.*, de Wet, *et al.*, 1985. *Proc. Natl. Acad. Sci. USA* **80**: 7870-7873) and plants (see *e.g.*, Ow, *et al.*, 1986. *Science* **234**: 856-859), as well as in insect (see *e.g.*, Jha, *et al.*, 1990. *FEBS Lett.* **274**: 24-26) and mammalian cells (see *e.g.*, de Wet, *et al.*, 1987. *Mol. Cell. Biol.* **7**: 725-7373; Keller, *et al.*, 1987. *Proc. Natl. Acad. Sci. USA* **82**: 3264-3268). In addition, a number of luciferase genes from the Jamaican click beetle, *Pyroplorus plagiophihalamus* (*Coleoptera*), have recently been cloned and partially characterized (see *e.g.*, Wood, *et al.*, 1989. *J. Biolumin. Chemilumin.* **4**: 289-301; Wood, *et al.*, 1989. *Science* **244**: 700-702). Distinct luciferases can sometimes produce light of different wavelengths, which may enable simultaneous monitoring of light emissions at different wavelengths. Accordingly, these aforementioned characteristics are unique, and add new dimensions with respect to the utilization of current reporter systems.

Firefly luciferase catalyzes bioluminescence in the presence of luciferin, adenosine 5'-triphosphate (ATP), magnesium ions, and oxygen, resulting in a quantum yield of 0.88 (see *e.g.*, McElroy and Selinger, 1960. *Arch. Biochem. Biophys.* **88**: 136-145). The firefly luciferase bioluminescent reaction can be utilized as an assay for the detection of ATP with a detection limit of approximately $1x10^{-13}$ M (see *e.g.*, Leach, 1981. *J. Appl. Biochem.* **3**: 473-517). In addition, the overall degree of sensitivity and convenience of the luciferase-mediated detection systems have created considerable interest in the development of firefly luciferase-based biosensors (see *e.g.*, Green and Kricka, 1984. *Talanta* **31**: 173-176; Blum, *et al.*, 1989. *J. Biolumin. Chemilumin.* **4**: 543-550).

Using the above-described enzymes, the sequence primer is exposed to a polymerase and a known dNTP. If the dNTP is incorporated onto the 3' end of the primer sequence, the dNTP is cleaved and a PPi molecule is liberated. The PPi is then converted to ATP with ATP sulfurylase.

Preferably, the ATP sulfurylase is present at a sufficiently high concentration that the conversion of PPi proceeds with first-order kinetics with respect to PPi. In the presence of luciferase, the ATP is hydrolyzed to generate a photon. The reaction preferably has a sufficient concentration of luciferase present within the reaction mixture such that the reaction, ATP $\rightarrow$ ADP + PO_4^{3-} + photon (light), proceeds with first-order kinetics with respect to ATP. The photon can be measured using methods and apparatuses described below.

For most applications it is desirable to use reagents free of contaminants like ATP and PPi . These contaminants may be removed by flowing the reagents through a precolumn containing apyrase and/-or pyrophosphatase bound to resin. Alternatively, the apyrase or pyrophosphatase can be bound to magnetic beads and used to remove contaminating ATP and PPi present in the reagents. In addition it is desirable to wash away diffusible sequencing reagents, *e.g.*, unincorporated dNTPs, with a wash buffer. Any wash buffer used in pyrophosphate sequencing can be used.

In some embodiments, the concentration of reactants in the sequencing reaction include 1 pmol DNA, 3 pmol polymerase, 40 pmol dNTP in 0.2 ml buffer. See Ronaghi, *et al.*, *Anal. Biochem.* 242: 84-89 (1996).

The sequencing reaction can be performed with each of four predetermined nucleotides, if desired. A "complete" cycle generally includes sequentially administering sequencing reagents for each of the nucleotides dATP, dGTP, dCTP and dTTP (or dUTP), in a predetermined order. Unincorporated dNTPs are washed away between each of the nucleotide additions. Alternatively, unincorporated dNTPs are degraded by apyrase (see below). The cycle is repeated as desired until the desired amount of sequence of the sequence product is obtained. In some embodiments, about 10-1000, 10-100, 10-75, 20-50, or about 30 nucleotides of sequence information is obtained from extension of one annealed sequencing primer.

Luciferase can hydrolyze dATP directly with concomitant release of a photon. This results in a false positive signal because the hydrolysis occurs independent of incorporation of the dATP into the extended sequencing primer. To avoid this problem, a dATP analog can be used which is incorporated into DNA, *i.e.*, it is a substrate for a DNA polymerase, but is not a substrate for luciferase. One such analog is α-thio-dATP. Thus, use of α-thio-dATP avoids the

spurious photon generation that can occur when dATP is hydrolyzed without being incorporated into a growing nucleic acid chain.

Typically, the PPi-based detection is calibrated by the measurement of the light released following the addition of control nucleotides to the sequencing reaction mixture immediately after the addition of the sequencing primer. This allows for normalization of the reaction conditions. Incorporation of two or more identical nucleotides in succession is revealed by a corresponding increase in the amount of light released. Thus, a two-fold increase in released light relative to control nucleotides reveals the incorporation of two successive dNTPs into the extended primer.

If desired, apyrase may be "washed" or "flowed" over the surface of the solid support so as to facilitate the degradation of any remaining, non-incorporated dNTPs within the sequencing reaction mixture. Upon treatment with apyrase, any remaining reactants are washed away in preparation for the following dNTP incubation and photon detection steps. Alternatively, the apyrase may be bound to the solid support.

When the support is planar, the pyrophosphate sequencing reactions preferably take place in a thin reaction chamber that includes one optically transparent solid support surface and an optically transparent cover. Sequencing reagents may then be delivered by flowing them across the surface of the substrate. When the support is not planar, the reagents may be delivered by dipping the solid support into baths of any given reagents.

When the support is in the form of a cavitated array, *e.g.*, in the termini of a fiber optic reactor array (FORA) or other array of microwells , suitable delivery methods for reagents include flowing and washing and also, *e.g.*, flowing, spraying, electrospraying, ink jet delivery, stamping, ultrasonic atomization (Sonotek Corp., Milton, NY) and rolling. Preferably, all reagent solutions contain 10-20% ethylene glycol to minimize evaporation. When spraying is used, reagents are delivered to the FORA surface in a homogeneous thin layer produced by industrial type spraying nozzles (Spraying Systems, Co., Wheaton, IL) or atomizers used in thin layer chromatography (TLC), such as CAMAG TLC Sprayer (Camag Scientific Inc., Wilmington, NC). These sprayers atomize reagents into aerosol spray particles in the size range of 0.3 to 10 μm.

Electrospray deposition (ESD) of protein and DNA solutions is currently used to generate ions for mass spectrometric analysis of these molecules. Deposition of charged electrospray products on certain areas of a FORA substrate under control of electrostatic forces is suggested. It was also demonstrated that the ES-deposited proteins and DNA retain their ability to specifically bind antibodies and matching DNA probes, respectively, enabling use of the ESD fabricated matrixes in Dot Immuno-Binding (DIB) and in DNA hybridization assays. (Morozov VN, Morozova TY: Electrospray deposition as a method for mass fabrication of mono- and multicomponent microarrays of biological and biologically active substances. Anal Chem 1999 Aug 1;71(15):3110-7)

Ink-jet delivery is applicable to protein solutions and other biomacromolecules, as documented in the literature (e.g. Roda A, Guardigli M, Russo C, Pasini P, Baraldini M., Protein microdeposition using a conventional ink-jet printer. Biotechniques 2000 Mar; 28(3): 492-6). It is also commercially available e.g. from MicroFab Technologies, Inc. (Plano, TX).

Reagent solutions can alternatively be delivered to the FORA surface by a method similar to lithography. Rollers (stamps; hydrophilic materials should be used) would be first covered with a reagent layer in reservoirs with dampening sponges and then rolled over (pressed against) the FORA surface.

Successive reagent delivery steps are preferably separated by wash steps. These washes can be performed, *e.g.*, using the above described methods, including high-flow sprayers or by a liquid flow over the FORA or microwell array surface.

In various embodiments, some components of the reaction are immobilized, while other components are provided in solution. For example, in some embodiments, the enzymes utilized in the pyrophosphate sequencing reaction (*e.g.,* sulfurylase, luciferase) may be immobilized if desired onto the solid support. Similarly, one or more or of the enzymes utilized in the pyrophosphate sequencing reaction, *e.g.,* sulfurylase, luciferase may be immobilized at the termini of a fiber optic reactor array. Other components of the reaction, *e.g.*, a polymerase (such as Klenow fragment), nucleic acid template, and nucleotides can be added by flowing, spraying, or rolling. In still further embodiments, one more of the reagents used in the sequencing reactions is delivered on beads.

In some embodiments, reagents are dispensed using an expandable, flexible membrane to dispense reagents and seal reactors on FORA surface during extension reactions. Reagents can be sprayed or rolled onto either the FORA surface or onto the flexible membrane. The flexible membrane could then be either rapidly expanded or physically moved into close proximity with the FORA thereby sealing the wells such that PPi would be unable to diffuse from well to well. Preferably, data acquisition takes place at a reasonable time after reaction initiation to allow maximal signal to generate.

A sequence in an extended anchor primer can also be identified using sequencing methods other than by detecting a sequence byproduct. For example, sequencing can be performed by measuring incorporation of labeled nucleotides or other nucleotide analogs. These methods can be used in conjunction with fluorescent or electrochemiluminescent-based methods.

Alternatively, sequence byproducts can be generated using dideoxynucleotides having a label on the 3' carbon. Preferably, the label can be cleaved to reveal a 3' hydroxyl group. In this method, addition of a given nucleotide is scored as positive or negative, and one base is determined at each trial. In this embodiment, solid phase enzymes are not required and multiple measurements can be made.

In another embodiment, the identity of the extended anchor primer product is determined using labeled deoxynucleotides. The labeled deoxynucleotides can be, *e.g.*, fluorescent nucleotides. Preferably the fluorescent nucleotides can be detected following laser-irradiation. Preferably, the fluorescent label is not stable for long periods of exposure. If desired, the fluorescent signal can be quenched, *e.g.*, photobleached, to return signal to background levels prior to addition of the next base. A preferred electrochemiluminescent label is ruthenium-tris-bi-pyridyl.

When luciferase is immobilized, it is preferably less than 50 μm from an anchored primer.

The photons generated by luciferase may be quantified using a variety of detection apparatuses, *e.g.*, a photomultiplier tube, charge-coupled device (CCD), CMOS, absorbance photometer, a luminometer, charge injection device (CID), or other solid state detector, as well as the apparatuses described herein. In a preferred embodiment, the quantitation of the emitted photons is accomplished by the use of a CCD camera fitted with a fused fiber optic bundle. In

another preferred embodiment, the quantitation of the emitted photons is accomplished by the use of a CCD camera fitted with a microchannel plate intensifier. A back-thinned CCD can be used to increase sensitivity. CCD detectors are described in, *e.g.*, Bronks, *et al.*, 1995. *Anal. Chem.* 65: 2750-2757.

An exemplary CCD system is a Spectral Instruments, Inc. (Tucson, AZ) Series 600 4-port camera with a Lockheed-Martin LM485 CCD chip and a 1-1 fiber optic connector (bundle) with 6-8 μm individual fiber diameters. This system has 4096x4096, or greater than 16 million pixels and has a quantum efficiency ranging from 10% to > 40%. Thus, depending on wavelength, as much as 40% of the photons imaged onto the CCD sensor are converted to detectable electrons.

Apparatuses for Sequencing Nucleic Acids

Also provided in the invention are apparatuses for sequencing nucleic acids. In some embodiments, the apparatuses include anchor primers attached to planar substrates. Nucleic acid sequence information can be detected using conventional optics or fiber-optic based systems attached to the planar substrate. In other embodiments, the apparatuses include anchor primers attached to the termini of fiber-optic arrays. In these embodiments, sequence information can be obtained directly from the termini of the fiber optic array.

Apparatus for sequencing nucleic acids

An apparatus for sequencing nucleic acids is illustrated in FIG. 2. The apparatus includes an inlet conduit 200 in communication with a detachable perfusion chamber 220. The inlet conduit 200 allows for entry of sequencing reagents via a plurality of tubes 202-212, which are each in communication with a plurality of sequencing dispensing reagent vessels 214-224.

Reagents are introduced through the conduit 200 into the perfusion chamber 220 using either a pressurized system or pumps to drive positive flow. Typically, the reagent flow rates are from 0.05 to 50 ml/minute (*e.g.*, 1 to 50 ml/minute) with volumes from 0.100 ml to continuous flow (for washing). Valves are under computer control to allow cycling of nucleotides and wash reagents. Sequencing reagents, *e.g.*, polymerase can be either pre-mixed with nucleotides or added in stream. A manifold brings all six tubes 202-212 together into one for feeding the perfusion chamber. Thus several reagent delivery ports allow access to the perfusion chamber.

For example, one of the ports may be utilized to allow the input of the aqueous sequencing reagents, while another port allows these reagents (and any reaction products) to be withdrawn from the perfusion chamber.

The perfusion chamber 200 contains a substrate to which a plurality of anchor primers have been attached. This can be a planar substrate containing one or more anchored primers in anchor pads formed at the termini of a bundled fiber optic array. The latter substrate surface is discussed in more detail below.

The perfusion chamber allows for a uniform, linear flow of the required sequencing reagents, in aqueous solution, over the amplified nucleic acids and allows for the rapid and complete exchange of these reagents. Thus, it is suitable for performing pyrophosphate-based sequencing reactions. The perfusion chamber can also be used to prepare the anchor primers and perform amplification reactions, *e.g.*, the RCA reactions described herein.

The solid support is optically linked to an imaging system 230, which includes a CCD system in association with conventional optics or a fiber optic bundle. In one embodiment the perfusion chamber substrate includes a fiber optic array wafer such that light generated near the aqueous interface is transmitted directly through the optical fibers to the exterior of the substrate or chamber. When the CCD system includes a fiber optic connector, imaging can be accomplished by placing the perfusion chamber substrate in direct contact with the connector. Alternatively, conventional optics can be used to image the light, *e.g.*, by using a 1-1 magnification high numerical aperture lens system, from the exterior of the fiber optic substrate directly onto the CCD sensor. When the substrate does not provide for fiber optic coupling, a lens system can also be used as described above, in which case either the substrate or the perfusion chamber cover is optically transparent. An exemplary CCD imaging system is described above.

The imaging system 230 is used to collect light from the reactors on the substrate surface. Light can be imaged, for example, onto a CCD using a high sensitivity low noise apparatus known in the art. For fiber-optic based imaging, it is preferable to incorporate the optical fibers directly into the cover slip or for a FORA to have the optical fibers that form the microwells also be the optical fibers that convey light to the detector.

The imaging system is linked to a computer control and data collection system 240. In general, any commonly available hardware and software package can be used. The computer control and data collection system is also linked to the conduit 200 to control reagent delivery.

An example of a perfusion chamber of the present invention is illustrated in FIG. 3. The perfusion chamber includes a sealed compartment with transparent upper and lower slide. It is designed to allow flow of solution over the surface of the substrate surface and to allow for fast exchange of reagents. Thus, it is suitable for carrying out, for example, the pyrophosphate sequencing reactions. . The shape and dimensions of the chamber can be adjusted to optimize reagent exchange to include bulk flow exchange, diffusive exchange, or both in either a laminar flow or a turbulent flow regime.

The perfusion chamber is preferably detached from the imaging system while it is being prepared and only placed on the imaging system when sequencing analysis is performed.

In one embodiment, the solid support (*i.e.*, a DNA chip or glass slide) is held in place by a metal or plastic housing, which may be assembled and disassembled to allow replacement of said solid support.

The lower side of the solid support of the perfusion chamber carries the reaction center array and, with a traditional optical-based focal system, a high numerical aperture objective lens is used to focus the image of the reaction center array onto the CCD imaging system.
The photons generated by the pyrophosphate sequencing reaction are captured by the CCD only if they pass through a focusing device (*e.g.*, an optical lens or optical fiber) and are focused upon a CCD element. However, the emitted photons will escape equally in all directions. In order to maximize their subsequent "capture" and quantitation when utilizing a planar array (*e.g.*, a DNA chip), it is preferable to collect the photons as close as possible to the point at which they are generated, *e.g.* immediately at the planar solid support. This is accomplished by either: (*i*) utilizing optical immersion oil between the cover slip and a traditional optical lens or optical

fiber bundle or, preferably, (*ii*) incorporating optical fibers directly into the cover slip itself. Similarly, when a thin, optically transparent planar surface is used, the optical fiber bundle can also be placed against its back surface, eliminating the need to "image" through the depth of the entire reaction/perfusion chamber.

Fiber optic substrate arrays with linked anchor primers

In some embodiments, the solid support is coupled to a bundle of optical fibers that are used to detect and transmit light generated by enzymatic processing of sequence reaction byproducts. The total number of optical fibers within the bundle may be varied so as to match the number of individual arrays utilized in the sequencing reaction. The number of optical fibers incorporated into the bundle is designed to match the CCD (*i.e.*, approximately 60 mm x 60mm) so as to allow 1:1 imaging. The desired number of optical fibers are initially fused into a bundle, the terminus of which is cut and polished so as to form a "wafer" of the required thickness (*e.g.*, 1.5 mm). The resulting optical fiber wafers possess similar handling properties to that of a plane of glass. The individual fibers can be any size diameter (*e.g.*, 3μm to 100 μm).

In some embodiments two fiber optic bundles are used: a first bundle is attached directly to the CCD sensor (the *fiber bundle* or *connector* or solid support) and a second bundle is used as the perfusion chamber substrate (the *wafer* or *substrate*). In this case the two are placed in direct contact, optionally with the use of optical coupling fluid, in order to image the reaction centers onto the CCD sensor. The overall sizes of the bundles are chosen so as to optimize the usable area of the CCD while maintaining desirable reagent (flow) characteristics in the perfusion chamber. Thus for a 4096 x 4096 pixel CCD array with 15 um pixels, the fiber bundle is chosen to be approximately 60 mm x 60 mm or to have a diameter of approximately 90 mm. The wafer could be slightly larger in order to maximize the use of the CCD area, or slightly smaller in order to match the format of a typical microscope slide—25 mm x 75 mm. The diameters of the individual fibers within the bundles are chosen so as to maximize the probability that a single reaction will be imaged onto a single CCD pixel, within the constraints of the state of the art. Exemplary diameters are 6-8 um for the fiber bundle and 6-50 um for the wafer, though any diameter in the range 3-100 um can be used. The fiber bundle is obtained commercially from the CCD camera manufacturer. The wafer can be obtained from Incom, Inc. (Charlton, MA) and is

cut and polished from a large fusion of fiber optics, typically being 2 mm thick, though possibly being 0.5 to 5 mm thick. The wafer has handling properties similar to a pane of glass or a glass microscope slide.

In other embodiments, the planar support is omitted and the anchor primers are linked directly to the termini of the optical fibers. Preferably, the anchor primers are attached to termini that are cavitated as shown schematically in FIG. 4. The termini are treated, *e.g.*, with acid, to form an indentation in the fiber optic material, wherein the indentation ranges in depth from approximately one-half the diameter of an individual optical fiber up to two to three times the diameter of the fiber.

Cavities can be introduced into the termini of the fibers by placing one side of the optical fiber wafer into an acid bath for a variable amount of time. The amount of time can vary depending upon the overall depth of the reaction cavity desired (see *e.g.*, Walt, *et al.*, 1996. *Anal. Chem.* 70: 1888). Several methods are known in the art for attaching molecules (and detecting the attached molecules) in the cavities etched in the ends of fiber optic bundles. See, *e.g.*, Michael, *et al.*, *Anal. Chem.* 70: 1242-1248 (1998); Ferguson, *et al.*, *Nature Biotechnology* 14: 1681-1684 (1996); Healey and Walt, *Anal. Chem.* 69: 2213-2216 (1997). A pattern of reactive sites can also be created in the microwell, using photolithographic techniques similar to those used in the generation of a pattern of reaction pads on a planar support. See, Healey, *et al.*, *Science* 269: 1078-1080 (1995); Munkholm and Walt, *Anal. Chem.* 58: 1427-1430 (1986), and Bronk, *et al.*, *Anal. Chem.* 67: 2750-2757 (1995).

The opposing side of the optical fiber wafer (*i.e.*, the non-etched side) is highly polished so as to allow optical-coupling (*e.g.*, by immersion oil or other optical coupling fluids) to a second, optical fiber bundle. This second optical fiber bundle exactly matches the diameter of the optical wafer containing the reaction chambers, and serve to act as a conduit for the transmission of the photons, generated by the pyrophosphate sequencing reaction, to its attached CCD imaging system or camera.

In one preferred embodiment, the fiber optic wafer is thoroughly cleaned, *e.g.* by serial washes in 15% H_2O_2/15%NH_4OH volume:volume in aqueous solution, then six deionized water rinses, then 0.5M EDTA, then six deionized water, then 15% H_2O_2/15%NH_4OH, then six deionized water (one-half hour incubations in each wash).

The surface of the fiber optic wafer is preferably coated to facilitate its use in the sequencing reactions. A coated surface is preferably optically transparent, allows for easy attachment of proteins and nucleic acids, and does not negatively affect the activity of immobilized proteins. In addition, the surface preferably minimizes non-specific absorption of macromolecules and increases the stability of linked macromolecules (*e.g.*, attached nucleic acids and proteins).

Suitable materials for coating the array include, *e.g.*, plastic (e.g. polystyrene). The plastic can be preferably spin-coated or sputtered (0.1 μm thickness). Other materials for coating the array include gold layers, *e.g.* 24 karat gold, 0.1 μm thickness, with adsorbed self-assembling monolayers of long chain thiol alkanes. Biotin is then coupled covalently to the surface and saturated with a biotin-binding protein (e.g. streptavidin).

Coating materials can additionally include those systems used to attach an anchor primer to a substrate. Organosilane reagents, which allow for direct covalent coupling of proteins via amino, sulfhydryl or carboxyl groups, can also be used to coat the array. Additional coating substances include photoreactive linkers, *e.g.* photobiotin, (Amos et al., "Biomaterial Surface Modification Using Photochemical Coupling Technology," in *Encyclopedic Handbook of Biomaterials and Bioengineering, Part A: Materials*, Wise et al. (eds.), New York, Marcel Dekker, pp. 895926, 1995).

Additional coating materials include hydrophilic polymer gels (polyacrylamide, polysaccharides), which preferably polymerize directly on the surface or polymer chains covalently attached post polymerization (Hjerten, J., J.Chromatogr. 347,191 (1985); Novotny, M., Anal. Chem. 62,2478 (1990), as well as pluronic polymers (triblock copolymers, e.g. PPO-PEO-PPO, also known as F-108), specifically adsorbed to either polystyrene or silanized glass surfaces (Ho et al., Langmuir 14:3889-94, 1998), as well as passively adsorbed layers of biotin-binding proteins.

In addition, any of the above materials can be derivatized with one or more functional groups, commonly known in the art for the immobilization of enzymes and nucleotides, *e.g.* metal chelating groups (e.g. nitrilo triacetic acid, iminodiacetic acid, pentadentate chelator), which will bind 6xHis-tagged proteins and nucleic acids.

In addition, surface coatings can be used that increase the number of available binding sites for subsequent treatments, *e.g.* attachment of enzymes (discussed later), beyond the theoretical binding capacity of a 2D surface.

In a preferred embodiment, the individual optical fibers utilized to generate the fused optical fiber bundle/wafer are larger in diameter (*i.e.*, 6 μm to 12 μm) than those utilized in the optical imaging system (*i.e.*, 3 μm). Thus, several of the optical imaging fibers can be utilized to image a single reaction site.

Mathematical analysis underlying optimization of the pyrophosphate sequencing reaction

While not wishing to be bound by theory, it is believed that optimization of reaction conditions can be performed using assumptions underlying the following analyses.

Solid-phase pyrophosphate sequencing was initially developed by combining a solid-phase technology and a sequencing-by-synthesis technique utilizing bioluminescence (see *e.g.*, Ronaghi, *et al.*, 1996. Real-time DNA sequencing using detection of pyrophosphate release. *Anal. Biochem.* **242**: 84-89). In the solid-phase methodology, an immobilized, primed DNA strand is incubated with DNA polymerase, ATP sulfurylase, and luciferase. By stepwise nucleotide addition with intermediate washing, the event of sequential polymerization can be followed. The signal-to-noise ratio was increased by the use of α-thio dATP in the system. This dATP analog is efficiently incorporated by DNA polymerase but does not serve as a substrate for luciferase. This reduces background bioluminescence and facilitates performance of the sequencing reaction in real-time. In these early studies, sequencing of a PCR product using streptavidin-coated magnetic beads as a solid support was presented. However, it was found that the loss of the beads during washing, which was performed between each nucleotide and enzyme addition, limited the technique to short sequences.

Currently, pyrophosphate sequencing methodologies have a reasonably well-established history for ascertaining the DNA sequence from many identical copies of a single DNA sequencing template (see *e.g.*, Ronaghi, *et al.*, 1996. Real-Time DNA Sequencing Using Detection of Pyrophosphate Release, *Anal. Biochem.* **242**: 84-89; Nyrén, *et al.*, Method of Sequencing DNA, patent WO9813523A1 (issued April 2, 1998; filed Sept. 26, 1997); Ronaghi,

et al., 1998. A Sequencing Method Based on Real-Time Pyrophosphate *Science* **281**: 363-365 (1998). Pyrophosphate (PPi)-producing reactions can be monitored by a very sensitive technique based on bioluminescence (see *e.g.*, Nyrén, *et al.*, 1996. pp. 466-496 (*Proc. 9th Inter. Symp. Biolumin. Chemilumin.*). These bioluminometric assays rely upon the detection of the PPi released in the different nucleic acid-modifying reactions. In these assays, the PPi which is generated is subsequently converted to ATP by ATP sulfurylase and the ATP production is continuously monitored by luciferase. For example, in polymerase-mediated reactions, the PPi is generated when a nucleotide is incorporated into a growing nucleic acid chain being synthesized by the polymerase. While generally, a DNA polymerase is utilized to generate PPi during a pyrophosphate sequencing reaction (see *e.g.*, Ronaghi, *et al.*, 1998. *Doctoral Dissertation,* The Royal Institute of Technology, Dept. of Biochemistry (Stockholm, Sweden)), it is also possible to use reverse transcriptase (see *e.g.*, Karamohamamed, *et al.*, 1996. pp. 319-329 (*Proc. 9th Inter. Symp. Biolumin. Chemilumin.*) or RNA polymerase (see *e.g.*, Karamohamamed, *et al.*, 1998. *BioTechniques* **24**: 302-306) to follow the polymerization event.

For example, a bioluminometric primer extension assay has been utilized to examine single nucleotide mismatches at the 3'-terminus (see *e.g.*, Nyrén, *et al.*, 1997. *Anal. Biochem.* **244**: 367-373). A phage promoter is typically attached onto at least one of the arbitrary primers and, following amplification, a transcriptional unit may be obtained which can then be subjected to stepwise extension by RNA polymerase. The transcription-mediated PPi-release can then be detected by a bioluminometric assay (*e.g.*, ATP sulfurylase-luciferase). By using this strategy, it is likely to be possible to sequence double-stranded DNA without any additional specific sequencing primer. In a series of "run-off" assays, the extension by T7 phage RNA polymerase has been examined and was found to be rather slow (see *e.g.*, Kwok, *et al.*, 1990. *Nucl. Acids Res.* **18**: 999-1005). The substitution of an α-thio nucleotide analogs for the subsequent, correct natural deoxynucleotide after the 3'-mismatch termini, could decrease the rate of polymerization by 5-fold to 13-fold. However, after incorporation of a few bases, the rate of DNA synthesis is comparable with the rate observed for a normal template/primer.

Single-base detection by this technique has been improved by incorporation of apyrase to the system, which catalyzes NTP hydrolysis and reduces the nucleotide concentration far below

the K_m of DNA polymerase, effectively removing dNTP from a preceding step before proceeding to addition of the subsequent dNTP. The above-described technique provides a rapid and real-time analysis for applications in the areas of mutation detection and single-nucleotide polymorphism (SNP) analysis.

The pyrophosphate sequencing system uses reactions catalyzed sequentially by several enzymes to monitor DNA synthesis. Enzyme properties such as stability, specificity, sensitivity, K_M and k_{CAT} are important for the optimal performance of the system. In the pyrophosphate sequencing system, the activity of the detection enzymes (*i.e.*, sulfurylase and luciferase) generally remain constant during the sequencing reaction, and are only very slightly inhibited by high amounts of products (see *e.g.*, Ronaghi, *et al.*, 1998. *Doctoral Dissertation,* The Royal Institute of Technology, Dept. of Biochemistry (Stockholm, Sweden)). Sulfurylase converts each PPi to ATP in approximately 2.0 seconds (see *e.g.*, Nyrén and Lundin, 1985. *Anal. Biochem.* **151**: 504-509). The reported reaction conditions for 1 pmol PPi in 0.2 ml buffer (5 nM) are 0.3 U/ml ATP sulfurylase (ATP:sulfate adenylyltransferase; Prod. No. A8957; Sigma Chemical Co., St. Louis, MO) and 5 μM APS (see *e.g.*, Ronaghi, *et al.*, 1996. Real-Time DNA Sequencing Using Detection of Pyrophosphate Release, *Anal. Biochem.* **242**: 84-89). The manufacturer's information (Sigma Chemical Co., St. Louis, MO) for sulfurylase reports an activity of 5-20 units per mg protein (*i.e.*, one unit will produce 1.0 μmole of ATP from APS and PPi per minute at pH 8.0 at 30 C), whereas the specific activity has been reported elsewhere as 140 units per mg (see Karamohamed, *et al.*, 1999. Purification, and Luminometric Analysis of Recombinant *Saccharomyces cerevisiae* MET3 Adenosine Triphosphate Sulfurylase Expressed in *Escherichia coli, Prot. Express. Purification* **15**: 381-388). Due to the fact that the reaction conditions utilized in the practice of the present invention are similar to those reaction conditions reported in the aforementioned reference, the sulfurylase concentration within the assay was estimated as 4.6 nM. The K_M values for sulfurylase are [APS] = 0.5 μM and [PPi] = 7 μM. The generation of light by luciferase takes place in less than 0.2 seconds. The most critical reactions are the DNA polymerization and the degradation of nucleotides. The value of constants characterizing the enzymes utilized in the pyrophosphate sequencing methodology are listed below for reference:

Enzyme	K_M (μM)	k_{CAT} (S^{-1})
Klenow	0.18 (dTTP)	0.92
T_7 DNA Polymerase	0.36 (dTTP)	0.52
ATP Sulfurylase	0.56 (APS); 7.0 (PPi)	38
Firefly Luciferase	20 (ATP)	0.015
Apyrase	120 (ATP); 260 (ADP)	500 (ATP)

The enzymes involved in these four reactions compete for the same substrates. Therefore, changes in substrate concentrations are coupled. The initial reaction is the binding of a dNTP to a polymerase/DNA complex for chain elongation. For this step to be rapid, the nucleotide triphosphate concentration must be above the K_M of the DNA polymerase. If the concentration of the nucleotide triphosphates is too high, however, lower fidelity of the polymerase may be observed (see *e.g.,* Cline, *et al.,* 1996. PCR fidelity of Pfu DNA polymerase and other thermostable DNA polymerases. *Nucl. Acids Res.* 24: 3546-3551). A suitable range of concentrations is established by the K_M for the misincorporation, which is usually much higher (see *e.g.,* Capson, *et al.,* 1992. Kinetic characterization of the polymerase and exonuclease activity of the gene 43 protein of bacteriophage T4. *Biochemistry* **31**: 10984-10994). Although a very high fidelity can be achieved by using polymerases with inherent exonuclease activity, their use also holds the disadvantage that primer degradation may occur.

Although the exonuclease activity of the Klenow fragment of DNA polymerase I (Klenow) is low, it has been demonstrated that the 3'-terminus of a primer may be degraded with longer incubations in the absence of nucleotide triphosphates (see *e.g.,* Ronaghi, *et al.,* 1998. *Doctoral Dissertation,* The Royal Institute of Technology, Dept. of Biochemistry (Stockholm, Sweden)). Fidelity is maintained without exonuclease activity because an induced-fit binding mechanism in the polymerization step provides a very efficient selectivity for the correct dNTP. Fidelities of $1x10^5$ to $1x10^6$ have been reported (see *e.g.,* Wong, *et al.,* 1991. An induced-fit kinetic mechanism for DNA replication fidelity. *Biochemistry* **30**: 526-537). In pyrophosphate sequencing, exonuclease-deficient (exo-) polymerases, such as exo-Klenow or Sequenase®, have been confirmed to have high fidelity.

Estimates for the spatial and temporal constraints on the pyrophosphate sequencing methodology of the present invention have been calculated, wherein the system possesses a 1 cm^2 area with height approximately 50 μm, for a total volume of 5 μl. With respect to temporal constraints, the molecular species participating in the cascade of reactions are initially defined, wherein:

N = the DNA attached to the surface

PPi = the pyrophosphate molecule released

ATP = the ATP generated from the pyrophosphate

L = the light released by luciferase

It is further specified that N(0) is the DNA with no nucleotides added, N(1) has 1 nucleotide added, N(2) has 2 nucleotides added, and so on. The pseudo-first-order rate constants which relate the concentrations of molecular species are:

$N(n) \rightarrow N(n+1) + PP_i$ k_N

$PPi \rightarrow ATP$ k_P

$ATP \rightarrow L$ k_A

In addition, the diffusion constants D_P for PPi and D_A for ATP must also be specified. These values may be estimated from the following exemplar diffusion constants for biomolecules in a dilute water solution (see Weisiger, 1997. Impact of Extracellular and Intracellular Diffusion on Hepatic Uptake Kinetics Department of Medicine and the Liver Center, University of California, San Francisco, California, USA, dickw@itsa.ucsf.edu, http://dickw.ucsf.edu/papers/goresky97/chapter.html).

Molecule	$D/10^{-5}$ cm^2/sec	Method	Original Reference
Albumin	0.066	lag time	1
Albumin	0.088	light scattering	2
Water	1.940	NMR	3

wherein, Original Reference 1 is: Longsworth, 1954. Temperature dependence of diffusion in aqueous solutions, *J. Phys. Chem.* **58**: 770-773; Original Reference 2 is: Gaigalas, *et al.,* 1992. Diffusion of bovine serum albumin in aqueous solutions, *J. Phys. Chem.* **96**: 2355-2359; and Original Reference 3 is: Cheng, 1993. Quantitation of non-Einstein diffusion behavior of water in biological tissues by proton NMR diffusion imaging: Synthetic image calculations, *Magnet. Reson. Imaging* **11**: 569-583.

In order to estimate the diffusion constant of PPi, the following exemplar values may be utilized (see *CRC Handbook of Chemistry and Physics*, 1983. (W.E. Weast. Ed.) CRC Press, Inc., Boca Raton, FL):

Molecule	$D/10^{-5}$ cm^2/sec	Molecular Weight/amu
sucrose	0.5226	342.30
mannitol	0.682	182.18
penta-erythritol	0.761	136.15
glycolamide	1.142	N/A
glycine	1.064	75.07

The molecular weight of PPi is 174 amu. Based upon the aforementioned exemplar values, a diffusion constant of approximately 0.7×10^{-5} cm^2/sec for PPi is expected.

Enzymes catalyzing the three pyrophosphate sequencing reactions are thought to approximate Michaelis-Menten kinetics (see *e.g.* Stryer, 1988. *Biochemistry*, W. H. Freeman and Company, New York), which may be described:

$K_M = [E][S]/[ES]$,

$$\text{velocity} = V_{max}[S] / (K_M + [S]),$$

$$V_{max} = k_{CAT}[E_T]$$

where [S] is the concentration of substrate, [E] is the concentration of free enzyme, [ES] is the concentration of the enzyme-substrate complex, and $[E_T]$ is the total concentration of enzyme = [E] + [ES].

It is preferable that the reaction times are at least as fast as the solution-phase pyrophosphate-based sequencing described in the literature. That rate that a substrate is converted into product is

$$-d[S]/dt = k_{CAT}[E_T][S]/(K_M + [S])$$

The effective concentration of substrate may be estimated from the size of a replicated DNA molecule, at most $(10\ \mu m)^3$ and the number of copies (approximately 10,000), yielding a concentration of approximately 17 nM. This is this is smaller than the K_M for the enzymes described previously, and therefore the rate can be estimated to be

$$-d[S]/dt = (k_{CAT}/K_M)[E_T][S].$$

Thus, with pseudo first-order kinetics, the rate constant for disappearance of substrate depends on k_{CAT} and K_M, which are constants for a given enzyme, and $[E_T]$. Using the same enzyme concentrations reported in the literature will therefore produce similar rates.

The first step in the pyrophosphate sequencing reaction (*i.e.*, incorporation of a new nucleotide and release of PPi) will now be examined in detail. The preferred reaction conditions are: 1 pmol DNA, 3 pmol polymerase, 40 pmol dNTP in 0.2 ml buffer. Under the aforementioned, preferred reaction conditions, the K_M for nucleotide incorporation for the Klenow fragment of DNA polymerase I is 0.2 μM and for Sequenase 2.0™ (US Biochemicals, Cleveland, OH) is 0.4 μM, and complete incorporation of 1 base is less than 0.2 sec (see *e.g.*, Ronaghi, *et al.*, 1996. Real-Time DNA Sequencing Using Detection of Pyrophosphate Release, *Anal. Biochem.* **242**: 84-89) with a polymerase concentration of 15 nM.

In a 5 μl reaction volume, there are a total of 10,000 anchor primers with 10,000 sequencing primer sites each, or $1x10^8$ total extension sites = 0.17 fmol. Results which have

been previously published in the literature suggest that polymerase should be present at 3-times abundance, or 0.5 fmol, within the reaction mixture. The final concentration of polymerase is then 0.1 nM. It should be noted that these reaction conditions are readily obtained in the practice of the present invention.

As previously stated, the time required for the nucleotide addition reaction is no greater than 0.2 sec per nucleotide. Hence, if the reaction is allowed to proceed for a total of T seconds, then nucleotide addition should be sufficiently rapid that stretches of up to (T/0.2) identical nucleotides should be completely filled-in by the action of the polymerase. As discussed previously, the rate-limiting step of the pyrophosphate sequencing reaction is the sulfurylase reaction, which requires a total of approximately 2 sec to convert one PPi to ATP. Accordingly, a total reaction time which allows completion of the sulfurylase reaction, should be sufficient to allow the polymerase to "fill-in" stretches of up to 10 identical nucleotides. In random DNA species, regions of 10 or more identical nucleotides have been demonstrated to occur with a per-nucleotide probability of approximately 4^{-10}, which is approximately $1x10^{-6}$. In the 10,000 sequences which are extended from anchor primers in a preferred embodiment of the present invention, each of which will be extended at least 30 nucleotides and preferably 100 nucleotides, it is expected that approximately one run of 10 identical nucleotides will be present. Thus, it may be concluded that runs of identical nucleotides should not pose a difficulty in the practice of the present invention.

The overall size of the resulting DNA molecule is, preferably, smaller than the size of the anchoring pads (*i.e.*, 10 μm) and must be smaller than the distance between the individual anchoring pads (*i.e.*, 100 μm). The radius of gyration of a single-stranded DNA concatemer with N total nucleotides may be mathematically-estimated by the following equation: radius = b $(N/N_0)^{0.6}$, where b is the persistence length and N_0 is the number of nucleotides per persistence length; the exponent 0.6 is characteristic of a self-avoiding walk (see *e.g.*, Doi, 1986. *The Theory of Polymer Dynamics* (Clarendon Press, New York); Flory, 1953. *Principles of Polymer Chemistry* (Cornell University Press, New York)). Using single-stranded DNA as an example, b is 4 nm and N_0 is 13.6 nucleotides. (see *e.g.*, Grosberg, 1994. *Statistical Physics of*

Macromolecules (AIP Press, New York)). Using 10,000 copies of a 100-mer, $N = 1x10^6$ and the radius of gyration is 3.3 μm.

The diffusion of PPi will now be discussed in detail. In the reaction conditions utilized in the present invention, $[PP_i]$ is approximately 0.17 fmol in 5 μl, or 0.03 nM, and [sulfurylase] is 4.6 nM as described previously. In the first 2 sec of the reaction, about 7 % (0.002 nM) of PPi is consumed by sulfurylase, using GEPASI simulation software (see Mendes, P. (1993) GEPASI: a software package for modelling the dynamics, steady states and control of biochemical and other systems. Comput. Appl. Biosci. 9, 563-571.). The parameters used in simulation were $K_M(PPi) =$ 7 μM, $k_{CAT} = 38\ s^{-1}$, and [sulfurylase] = 4.6 nM. Therefore, it may be concluded that at least 93% of PPi molecules may diffuse away before being converted to ATP during the 2 sec reaction time.

The mean time for each PPi to react is $1/k_p = 2$ seconds. The mean square distance it diffuses in each direction is approximately $2D_p/k_p$, or $2.8x10^3\ \mu m^2$. The RMS distance in each direction is 53 μm. This value indicates that each of the individual anchor primers must be more than 50 μm apart, or PPi which is released from one anchor could diffuse to the next, and be detected.

Another method which may be used to explain the aforementioned phenomenon is to estimate the amount of PPi over a first anchor pad that was generated at said first anchor pad relative to the amount of PPi that was generated at a second anchor pad and subsequently diffused over to the location of said first anchor pad. When these two quantities approach each other in magnitude, it becomes difficult to distinguish the "true" signal from that of the background. This may be mathematically-described by defining **a** as the radius of an anchor pad and $1/b^2$ as the density of an anchor pad. Based upon previously published data, **a** is approximately equal to 10 μm and b is approximately equal to 100 μm. The amount of PPi which is present over said first anchor pad may be described by: $\exp(-k_pt)[1 - \exp(-\mathbf{a}^2/2D_pt)]$ and the amount of PPi present over the second anchor pads may be mathematically approximated by: $(1/3)\exp(-k_pt)[pa^2/b^2]\exp(-b^2/2D_pt)$. The prefactor 1/3 assumes that ¼ of the DNA sequences will incorporate 1 nucleotide, ¼ of these will then incorporate a second nucleotide, etc., and thus the

sum of the series is 1/3. The amounts of PPi over the first and second anchor pads become similar in magnitude when $2D_pt$ is approximately equal to b^2, thus indicating that the RMS distance a molecule diffuses is equal to the distance between adjacent anchor pads. In accord, based upon the assay conditions utilized in the practice of the present invention, the anchor pads must be placed no closer than approximately 50 μm apart, and preferable are at least 3-times further apart (*i.e.*, 150 μm).

Although the aforementioned findings set a limit on the surface density of anchor pads, it is possible to decrease the distance requirements, while concomitantly increasing the overall surface density of the anchor pads, by the use of a number of different approaches. One approach is to detect only the early light, although this has the disadvantage of losing signal, particularly from DNA sequences which possess a number of contiguous, identical nucleotides.

A second approach to decrease the distance between anchor pads is to increase the concentration of sulfurylase in the reaction mixture. The reaction rate k_p is directly proportional to the sulfurylase concentration, and the diffusion distance scales as $k_p^{-1/2}$. Therefore, if the sulfurylase enzyme concentration is increased by a factor of 4-times, the distance between individual anchor pads may be concomitantly reduced by a factor of 2-times.

A third approach is to increase the effective concentration of sulfurylase (which will also work for other enzymes described herein) by binding the enzyme to the surface of the anchor pads. The anchor pad can be approximated as one wall of a cubic surface enclosing a sequencing reaction center. Assuming a 10 μm x 10 μm surface for the pad, the number of molecules bound to the pad to produce a concentration of a 1 μM is approximately 600,000 molecules.

The sulfurylase concentration in the assay is estimated as 5nM. The number of bound molecules to reach this effective concentration is about 3000 molecules. Thus, by binding more enzyme molecules, a greater effective concentration will be attained. For example, 10,000 molecules could be bound per anchor pad.

As previously estimated, each sulfurylase molecule occupies a total area of 65 nm^2 on a surface. Accordingly, anchoring a total of 10,000 sulfurylase enzyme molecules on a surface (*i.e.*, so as to equal the 10,000 PPi released) would require 1.7 μm^2. This value is only

approximately 2% of the available surface area on a 10 μm x 10 μm anchor pad. Hence, the concentration of the enzyme may be readily increased to a much higher value.

A fourth approach to allow a decrease in the distance between individual anchor pads, is to utilize one or more agents to increase the viscosity of the aqueous-based, pyrophosphate sequencing reagents (*e.g.*, glycerol, polyethylene glycol (PEG), and the like) so as to markedly increase the time it takes for the PPi to diffuse. However, these agents will also concomitantly increase the diffusion time for other non-immobilized components within the sequencing reaction, thus slowing the overall reaction kinetics. Additionally, the use of these agents may also function to chemically-interfere with the sequencing reaction itself.

A fifth, and preferred, methodology to allow a decrease in the distance between individual anchor pads, is to conduct the pyrophosphate sequencing reaction in a spatial-geometry which physically-prevents the released PPi from diffusing laterally. For example, uniform cavities or microwells, such as those generated by acid-etching the termini of optical fiber bundles, may be utilized to prevent such lateral diffusion of PPi (see Michael, *et al.*, 1998. Randomly Ordered Addressable High-Density Optical Sensor Arrays, *Anal. Chem.* **70**: 1242-1248). In this embodiment, the important variable involves the total diffusion time for the PPi to exit a cavity of height **h**, wherein **h** is the depth of the etched cavity. This diffusion time may be calculated utilizing the equation: $2D_pt = \mathbf{h}^2$. By use of the preferred pyrophosphate sequencing reaction conditions of the present invention in the aforementioned calculations, it may be demonstrated that a cavity 50 μm in depth would be required for the sequencing reaction to proceed to completion before complete diffusion of the PPi from said cavity. Moreover, this type of geometry has the additional advantage of concomitantly reducing background signal from the PPi released from adjacent anchor pads.

Additionally, to prevent background generated by diffusion of PPi from one pad to another, the region of substrate between the pads can be coated with immobilized phosphatase.

Subsequently, once ATP has been formed by use of the preferred reaction conditions of the present invention, the reaction time, $1/k_A$, has been shown to be 0.2 seconds. Because this reaction time is much lower than the time which the PPi is free to diffuse, it does not

significantly alter any of the aforementioned conclusions regarding the assay geometry and conditions utilized in the present invention.

In order to mitigate the generation of background light, it is preferable to "localize" (*e.g.*, by anchoring or binding) the luciferase in the region of the DNA sequencing templates. It is most preferable to localize the luciferase to a region that is delineated by the distance a PPi molecule can diffuse before it forms ATP. Methods for binding luciferase to a solid support matrix are well-known in the literature (see *e.g.*, Wang, *et al.*, 1997. Specific Immobilization of Firefly Luciferase through a Biotin Carboxyl Carrier Protein Domain, *Analytical Biochem.* **246**: 133-139). Thus, for a 2 second diffusion time, the luciferase is anchored within a 50 μm distance of the DNA strand. It should be noted, however, that it would be preferable to decrease the diffusion time and thus to further limit the surface area which is required for luciferase binding.

Additionally, to prevent background generated by diffusion of ATP from one pad to another, the region of substrate between the pads can be coated with immobilized ATPase, especially one that hydrolyzes ATP to ADP, *e.g.* alkaline phosphatase.

In order to determine the concentration of luciferase which it is necessary to bind, previously published conditions were utilized in which luciferase is used at a concentration which gives a response of 200 mV for 0.1 μm ATP (see Ronaghi, *et al.*, 1996. Real-Time DNA Sequencing Using Detection of Pyrophosphate Release, *Analytical Biochem.* **242**: 84-89). More specifically, it is known from the literature that, in a 0.2 ml reaction volume, 2 ng of luciferase gives a response of 10 mV for 0.1 μM ATP (see Karamohamed and Nyrén, 1999. Real-Time Detection and Quantification of Adenosine Triphosphate Sulfurylase Activity by a Bioluminometric Approach, *Analytical Biochem.* **271**: 81-85). Accordingly, a concentration of 20 ng of luciferase within a 0.2 ml total reaction volume would be required to reproduce these previously-published literature conditions. In the volume of a 10μm cube around each of the individual anchor pads of the present invention, a luciferase concentration of $1x10^{-16}$ grams would be required, and based upon the 71 kD molecular weight of luciferase, this concentration would be equivalent to approximately 1000 luciferase molecules. As previously stated, the surface area of luciferase has been computed at 50 nm^2. Thus, assuming the luciferase molecules were biotinylated and bound to the anchor pad, 1000 molecules would occupy a total area of 0.05

μm^2. From these calculations it becomes readily apparent that a plethora of luciferase molecules may be bound to the anchor pad, as the area of each anchor pad area is 100 μm^2.

Again, based upon previously published results in the literature, each nucleotide takes approximately 3 seconds to sequence (*i.e.*, 0.2 second to add a nucleotide; 2 seconds to make ATP; 0.2 seconds to get bioluminescence). Accordingly, a cycle time of approximately 60 seconds per nucleotide is reasonable, requiring approximately 30 minutes per experiment to generate 30 nucleotides of information per sequencing template.

In an alternative embodiment to the aforementioned sequencing methodology (*i.e.*, polymerase → PPi → sulfurylase → ATP → luciferase→← light), a polymerase may be developed (*e.g.*, through the use of protein fusion and the like) which possesses the ability to generate light when it incorporates a nucleotide into a growing DNA chain. In yet another alternative embodiment, a sensor may be developed which directly measures the production of PPi in the sequencing reaction. As the production of PPi changes the electric potential of the surrounding buffer, this change could be measured and calibrated to quantify the concentration of PPi produced.

As previously discussed, the polymerase-mediated incorporation of dNTPs into the nucleotide sequence in the pyrophosphate sequencing reaction causes the release of a photon (*i.e.*, light). The photons generated by the pyrophosphate sequencing reaction may subsequently be "captured" and quantified by a variety of methodologies including, but not limited to: a photomultiplier tube, CCD, absorbance photometer, a luminometer, and the like.

The photons generated by the pyrophosphate sequencing reaction are captured by the CCD. The efficiency of light capture increases if they pass through a focusing device (*e.g.*, an optical lens or optical fiber) and are focused upon a CCD element. The fraction of these photons which are captured may be estimated by the following calculations. First, it is assumed that the lens that focuses the emitted photons is at a distance **r** from the surface of the solid surface (*i.e.*, DNA chip or etched fiber optic well), where **r** = 1 cm, and that the photons must pass through a region of diameter **b** (area = $\pi \mathbf{b}^2/4$) so as to be focused upon the array element, where **b** = 100 μ m. (This produces an optical system with numerical aperture of approximately 0.01 in air.) It should also be noted that the emitted photons should escape equally in all directions. At distance

r, the photons are dispersed over an area of which is equal to $4\pi\mathbf{r}^2$. Thus, the fraction of photons which pass through the lens is described by: $(1/2)[1 - (1 + \mathbf{b}^2/4\mathbf{r}^2)^{-1/2}]$. When the value of **r** is much larger than that of **b**, the fraction which pass through the lens may then be described by: $\mathbf{b}^2/16\mathbf{r}^2$. For the aforementioned values of **r** and **b**, this fraction of photons is $6x10^{-6}$. Note that the fraction of captured photons increases as **b** increases or **r** decreases (*i.e.* as the numerical aperture of the imaging system increases). Use of FORA in which the microwells are etched into the termini of optical fibers, which then also serve to focus the light onto a CCD, greatly increases the numerical aperture from the example given above, with the numerical aperture of many fiber optics being in the range of 0.7. For each nucleotide addition, it is expected that approximately 10,000 PPi molecules will be generated and, if all are converted by sulfurylase and luciferase, these PPi will result in the emission of approximately $1x10^4$ photons. In order to maximize their subsequent "capture" and quantitation when utilizing a planar array (*e.g.*, a DNA chip), it is preferable to collect the photons immediately at the planar solid support (*e.g.*, the cover slip). This may be accomplished by either: (*i*) utilizing optical immersion oil between the cover slip and a traditional optical lens or optical fiber bundle or, preferably, (*ii*) incorporating optical fibers directly into the cover slip itself. Performing the previously described calculations (where in this case, **b** = 100 μm and **r** = 50 μm), the fraction collected is found to be 0.15, which equates to the capture of approximately $1x10^3$ photons. This value would be sufficient to provide an adequate signal.

The following examples are meant to illustrate, not limit, the invention.

Example 1. Construction of Anchor Primers Linked to a Cavitated Terminus Fiber Optic Array

The termini of a thin wafer fiber optic array are cavitated by inserting the termini into acid as described by Healey et al., *Anal. Chem.* 69: 2213-2216 (1997).

A thin layer of a photoactivatable biotin analog is dried onto the cavitated surface as described Hengsakul and Cass (*Bioconjugate Chem.* 7: 249-254, 1996) and exposed to white light through a mask to create defined pads, or areas of active biotin. Next, avidin is added and allowed to bind to the biotin. Biotinylated oligonucleotides are then added. The avidin has free

biotin binding sites that can anchor biotinylated oligonucleotides through a biotin-avidin-biotin link.

The pads are approximately 10 μm on a side with a 100 μm spacing. Oligonucleotides are added so that approximately 37% of the pads include one anchored primer. On a 1 cm^2 surface are deposited 10,000 pads, yielding approximately 3700 pads with a single anchor primer.

Example 2. Annealing and Amplification of Members of a Circular Nucleic Acid Library

A library of open circle library templates is prepared from a population of nucleic acids suspected of containing a single nucleotide polymorphism on a 70 bp Sau3A1-MspI fragment. The templates include adapters that are complementary to the anchor primer, a region complementary to a sequencing primer, and an insert sequence that is to be characterized. The library is generated using Sau3A1 and MspI to digest the genomic DNA. Inserts approximately 65-75nucleotides are selected and ligated to adapter oligonucleotides 12 nucleotides in length. The adapter oligonucleotides have sequences complementary to sequences to an anchor primers linked to a substrate surface as described in Example 1.

The library is annealed to the array of anchor primers. A DNA polymerase is added, along with dNTPs, and rolling circle replication is used to extend the anchor primer. The result is a single DNA strand, still anchored to the solid support, that is a concatenation of multiple copies of the circular template. 10,000 or more copies of circular templates in the hundred nucleotide size range.

Example 3. Sequence Analysis of Nucleic Acid Linked to the Terminus of a Fiber Optic Substrate

The fiber optic array wafer containing amplified nucleic acids as described in Example 2 is placed in a perfusion chamber and attached to a bundle of fiber optic arrays, which are themselves linked to a 16 million pixel CCD cameras. A sequencing primer is delivered into the perfusion chamber and allowed to anneal to the amplified sequences. Then sulfurylase, apyrase, and luciferase are attached to the cavitated substrate using biotin-avidin.

The sequencing primer primes DNA synthesis extending into the insert suspected of having a polymorphism, as shown in FIG.1. The sequencing primer is first extended by delivering into the perfusion chamber, in succession, a wash solution, a DNA polymerase, and one of dTTP, dGTP, dCTP, or α thio dATP (a dATP analog). The sulfurylase, luciferase, and apyrase, attached to the termini convert any PPi liberated as part of the sequencing reaction to detectable light. The apyrase present degrades any unreacted dNTP. Light is typically allowed to collect for 3 seconds (although 1-100, *e.g.*, 2-10 seconds is also suitable) by a CCD camera linked to the fiber imaging bundle, after which additional wash solution is added to the perfusion chamber to remove excess nucleotides and byproducts. The next nucleotide is then added, along with polymerase, thereby repeating the cycle.

During the wash the collected light image is transferred from the CCD camera to a computer. Light emission is analyzed by the computer and used to determine whether the corresponding dNTP has been incorporated into the extended sequence primer. Addition of dNTPs and pyrophosphate sequencing reagents is repeated until the sequence of the insert region containing the suspected polymorphism is obtained.

Example 4. Sequence Analysis of a Tandem Repeat Template Generated Using Rolling Circle Amplification

A primer having the sequence 5'-gAC CTC ACA CgA Tgg CTg CAg CTT – 3' (SEQ ID NO:2) was annealed to a 88 nucleotide template molecule having the sequence 5'-TCg TgT gAg gTC TCA gCA TCT TAT gTA TAT TTA CTT CTA TTC TCA gTT gCC TAA gCT gCA gCC A-3' (SEQ ID NO:8). Annealing of the template to the primer resulted in juxtaposition of the 5' and 3' ands of the template molecule. The annealed template was exposed to ligase, which resulted in ligation of the 5' and 3' ends of the template to generate a circular molecule.

The annealed primer was extended using Klenow fragment and nucleotides in rolling circle amplification for 12 hours at 37 °C. The product was purified using the SPRI technique (Seradyn, Indianapolis, IN). Rolling circle amplification resulted in formation of tandem repeats of a sequence complementary to the circular template sequence.

The tandem repeat product in the extended sequence was identified by annealing a sequencing primer having the sequence 5'-AAgCTgCAgCCATCgTgTgAgg-3' (SEQ ID NO:8)

and subjecting the annealed primer to 40 alternating cycles of 95 °C, 1 minute, 20 seconds,60 °C using ET terminator chemistry (Amersham-Pharmacia) in the presence of 1M betaine.

The sequencing product was then diluted to 1/5 volume and purified on a G-50 Sephadex column prior to injection into a MegaBACE sequencing system with linear polyacrylamide (Amersham-Pharmacia).

An electropherogram of the sequencing analysis is shown in FIG. 5. The tracing demonstrates that multiple copies of the 88 bp circular template molecule are generated tandemly, and that these copies can be detected in a DNA sequencing reaction.

Other Embodiments

It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

WHAT IS CLAIMED IS:

1. A method for sequencing a nucleic acid, the method comprising:

providing one or more nucleic acid anchor primers;

providing a plurality of single-stranded circular nucleic acid templates;

annealing an effective amount of the nucleic acid anchor primer to at least one of the single-stranded circular templates to yield a primed anchor primer-circular template complex;

combining the primed anchor primer-circular template complex with a polymerase to form an extended anchor primer covalently linked to multiple copies of a nucleic acid complementary to the circular nucleic acid template;

annealing an effective amount of a sequencing primer to one or more copies of said covalently linked complementary nucleic acid;

extending the sequencing primer with a polymerase and a predetermined nucleotide triphosphate to yield a sequencing product and, if the predetermined nucleotide triphosphate is incorporated onto the 3' end of said sequencing primer, a sequencing reaction byproduct; and

identifying the sequencing reaction byproduct, thereby determining the sequence of the nucleic acid.

2. The method of claim 1, wherein said anchor primer is linked to a solid support.

3. The method of claim 2, wherein said anchor primer is linked to the solid support prior to formation of said extended anchor primer.

4. The method of claim 2, wherein said anchor primer is linked to the solid support after formation of said extended anchor primer.

5. The method of claim 2, wherein said anchor primer is linked to the solid support during formation of said extended anchor primer.

6. The method of claim 1, wherein the circular nucleic acid template is single-stranded DNA.

7. The method of claim 1, wherein the circular nucleic acid template is an open circle nucleic acid or a closed-circle nucleic acid.

8. The method of claim 1, wherein the circular nucleic acid template is genomic DNA or cDNA.

9. The method of claim 1, wherein the circular nucleic acid is 10-200 nucleotides in length.

10. The method of claim 1, wherein the circular nucleic acid is 20-1000 nucleotides in length.

11. The method of claim 1, wherein the circular nucleic acid is 40-500 nucleotides in length.

12. The method of claim 1, wherein the primed circular template is extended by rolling circle amplification to yield a single-stranded concatamer of the annealed circular nucleic acid template.

13 The method of claim 10, wherein synthesis is carried out with biotin-conjugated nucleotides.

14. The method of claim 10, further comprising:

annealing a reverse primer to the single-stranded concatamer to yield a primed concatamer template, and

combining the primed concatamer template with a polymerase and nucleotide triphosphates to generate multiple copies of the concatamer template.

15. The method of claim 1, wherein the sequencing byproduct is pyrophosphate.

16. The method of claim 15, wherein the pyrophosphate is detected by contacting the sequencing byproduct with ATP sulfurylase or nicotinamide-mononucleotide adenylyl transferase under conditions sufficient to form ATP.

17. The method of claim 15, wherein the pyrophosphate is detected by contacting the sequencing byproduct with nicotinamide-mononucleotide adenylyl transferase under conditions sufficient to form ATP.

18. The method of claim 16, wherein the sulfurylase is a thermostable sulfurylase.

19. The method of claim 17, wherein the adenylyl transferase is a thermostable adenylase.

20. The method of claim 16, wherein the ATP is detected by luciferase.

21. The method of claim 20, wherein the luciferase is a thermostable luciferase.

22. The method of claim 15, wherein at least one of the enzymes for detection of pyrophosphate is immobilized on the substrate.

23. The method of claim 15, further comprising apyrase.

24. The method of claim 15, further comprising washing the sequencing product with a wash buffer.

25. The method of claim 24, wherein the wash buffer includes apyrase.

26. The method of claim 1, wherein the anchor primer sequence includes a biotin group.

27. The method of claim 26, wherein the biotin group on the anchor primer is linked to a biotin-binding protein on the solid support.

28. The method of claim 1, wherein one or more of the anchor primers is linked to a polysaccharide.

29. The method of claim 28, wherein one or more of the anchor primers is linked to a plurality of avidin moieties.

30. The method of claim 28, wherein one or more of the anchor primers is linked to a hexahistidine tag and the polysaccharide chain includes a nitrilotriacetic acid (NTA) moiety.

31. The method of claim 28, wherein the polysaccharide is linked to the solid support.

32. The method of claim 15, wherein the ATP sulfurylase or nicotinamide-mononucleotide adenylyl transferase is linked to a polysaccharide.

33. The method of claim 20, wherein the luciferase is linked to a polysaccharide.

34 The method of claim 32, wherein one or more of the enzymes carry biotin groups and the polysaccharide chain possesses a plurality of avidin moieties.

35 The method of claim 32, wherein one or more of the enzymes carry hexahistidine and the polysaccharide chain possesses a plurality of NTA.

36. The method of claim 32, wherein the polysaccharide chain is linked to the solid support.

37. The method of claim 1, wherein the solid support includes at least one optical fiber.

38. The method of claim 1, wherein the sequencing primer is extended in the presence of a dATP analog.

39. The method of claim 38, wherein the dATP analog is α thio dATP.

40. The method of claim 1, wherein the solid substrate includes two or more anchor primers separated by approximately 10 μm to approximately 200 μm.

41. The method of claim 1, wherein the solid substrate includes two or more anchor primers separated by approximately 50 μm to approximately 150 μm.

42. The method of claim 40, wherein the solid substrate includes two or more anchor primers separated by approximately 100 μm to approximately 150 μm.

43. The method of claim 1, wherein the solid support matrix comprises a plurality of anchor pads that are linked to the solid support.

44. The method of claim 1, wherein the solid support matrix comprises a plurality of anchor pads that are covalently linked to the solid support.

45. The method of claim 30, wherein the surface area of each anchor pad is approximately 10 μm^2.

46. The method of claim 30, wherein each pad is separated from one another by a distance ranging from approximately 50 μm to approximately 150 μm.

47. The method of claim 1, wherein said method comprises sequencing at least 100 nucleic acids.

48. The method of claim 1, wherein said method comprises sequencing at least 1000 nucleic acids.

49. The method of claim 1, wherein said method comprises sequencing at least 10,000 nucleic acids.

50. A method for analyzing a mixture of polymers, the method comprising:

providing a plurality of at least 1,000 independent reaction sites attached to a solid substrate, wherein said independent reaction sites comprise a reaction center and said independent reactions sites are separated by distances that eliminate diffusion of reactants and products between sites;

contacting said plurality of reaction sites with a fluid comprising said mixture of polymers; and

separating said polymers from the fluid and from each other by attachment to the reaction sites, thereby analyzing said mixture of biopolymers.

51. The method of claim 50, wherein said polymer is a biopolymer.

52. The method of claim 50, wherein said polymer is DNA, RNA, or a polypeptide.

53. The method of claim 50, wherein said medium is a low diffusivity medium.

54. The method of claim 50, wherein said biopolymers are separated by electrophoresis.

55. The method of claim 50, wherein said reaction sites are disposed on a planar surface.

56. A substrate for analyzing a nucleic acid, the substrate comprising:

a cavitated fiber optic surface; and

a nucleic acid sequence linked to the fiber optic surface.

57. The substrate of claim 56, wherein the nucleic acid sequence is an anchor primer.

58. The substrate of claim 56, wherein the substrate comprises a plurality of fiber optic surfaces.

59. The substrate of claim 58, wherein the fiber optic surface includes two or more anchoring primers separated by approximately 10 μm to approximately 200 μm.

60. The substrate of claim 58, wherein the fiber optic surface includes two or more anchoring primers separated by approximately 100 μm to approximately 150 μm.

61. The substrate of claim 58, wherein the fiber optic surface includes two or more anchoring primers separated by approximately 150 μm.

62. The substrate of claim 58, wherein the fiber optic surface includes two or more anchor pads separated by approximately 100 μm to approximately 150 μm.

63. The substrate of claim 62, wherein the surface area of each pad is approximately 10 μm^2.

64. A substrate with a cavitated surface comprising 10^3 or more groups of oligonucleotides attached to the surface in discrete known regions, the 10^3 or more groups of oligonucleotides occupying a total area of less than 1 cm^2 on said substrate, said groups of oligonucleotides having different nucleotide sequences.

65. The substrate of claim 64, wherein said substrate comprises 10^4 or more different groups of sequences in discrete known regions.

66. The substrate of claim 64, wherein said substrate comprises 10^5 or more different groups of oligonucleotides with known sequences in discrete known regions.

67. The substrate of claim 64, wherein the groups of oligonucleotides are attached to the surface by a linker.

68. The substrate of claim 64, wherein the groups of oligonucleotides are covalently attached to the surface.

69. An array of more than 1,000 different groups of oligonucleotide molecules with known sequences covalently coupled to a surface of a cavitated substrate, said groups of oligonucleotide molecules each in discrete known regions and differing from other groups of oligonucleotide molecules in monomer sequence, each of said discrete known regions being an area of less than about 0.01 cm^2 and each discrete known region comprising oligonucleotides of known sequence, said different groups occupying a total area of less than 1 cm^2.

70. The array of claim 69, wherein said area is less than 10,000 $microns^2$.

71. The array of claim 69, wherein said array is made by the process of:

exposing a first region of said substrate to light to remove a photoremovable group from photoprotected first region, and not exposing a second region of said surface to light; chemical activation of deprotected areas, covalently coupling a first nucleotide to said activated areas, applying a new mask to deprotect areas not deprotected in the first step, repeating the nucleotide coupling step thus covalently coupling a second nucleotide to said region exposed to light; and

repeating said steps of exposing said substrate to light and covalently coupling nucleotides until more than 500 different groups of nucleotides are formed on said surface.

72. The array of claim 69, wherein said array comprises more than 10,000 groups of oligonucleotides of known sequences.

73. An apparatus for analyzing a nucleic acid sequence, the apparatus comprising:

a reagent delivery chamber, wherein the chamber includes a substrate with immobilized nucleic acids;

a conduit in communication with the reagent delivery chamber;

an imaging system in communication with the reagent delivery chamber; and

a data collection system in communication with the imaging system.

74. The apparatus of claim 73, wherein the substrate is a planar substrate.

75. The apparatus of claim 73, wherein the substrate is a cavitated planar substrate.

76. The apparatus of claim 73, wherein the imaging system is a fiber optic system.

77. The apparatus of claim 73, wherein the substrate comprises

a cavitated fiber optic surface in communication with said imaging system; and

a nucleic acid sequence linked to the fiber optic surface.

78. The apparatus of claim 77, wherein the substrate comprises a plurality of fiber optic surfaces, said fiber optic surfaces being in communication with said imaging system.

79. The apparatus of claim 78, wherein the fiber optic surfaces include two or more anchoring primers separated by approximately 100 μm to approximately 150μm.

80. The apparatus of claim 78, wherein the fiber optic surfaces include two or more anchoring primers separated by approximately 150 μm.

81. The apparatus of claim 78, wherein the fiber optic surfaces include two or more anchor pads separated by approximately 100 μm to approximately 150 μm.

82. The apparatus of claim 78, wherein the surface area of each pad is approximately 5 μm^2 to approximately 10,000 μm^2.

83. The apparatus of claim 78, wherein the surface area of each pad is approximately 10 μm^2.

84. An apparatus for processing a plurality of analytes, the apparatus comprising:

a flow chamber having disposed therein a substrate comprising a plurality of cavitated surfaces, said cavitated surfaces having disposed thereon nucleic acid molecules;

fluid means for delivering processing reagents from one or more reservoirs to the flow chamber so that the analytes anchored to the plurality of microparticles are exposed to the reagents; and

detection means for detecting a sequence of optical signals from each microparticle of the plurality, each optical signal of the sequence being indicative of an interaction between a processing reagent and the analyte anchored thereto, wherein said detection means is in communication with the cavitated surfaces.

85. The apparatus of claim 85, wherein said detection means further comprises signal tracking means for correlating said optical signals from each of said microparticles in each of said digital images to form for each said microparticle of said plurality a sequence of said optical signals.

86. The apparatus of claim 87, wherein said signal tracking means is a CCD camera.

87. The apparatus of claim 86, wherein said analyte is DNA.

ABSTRACT

Disclosed herein are methods and apparatuses for sequencing a nucleic acid. These methods permit a very large number of independent sequencing reactions to be arrayed in parallel, permitting simultaneous sequencing of a very large number (>10,000) of different oligonucleotides.

TRADOCS:1448303.4(v1$n04!.DOC)

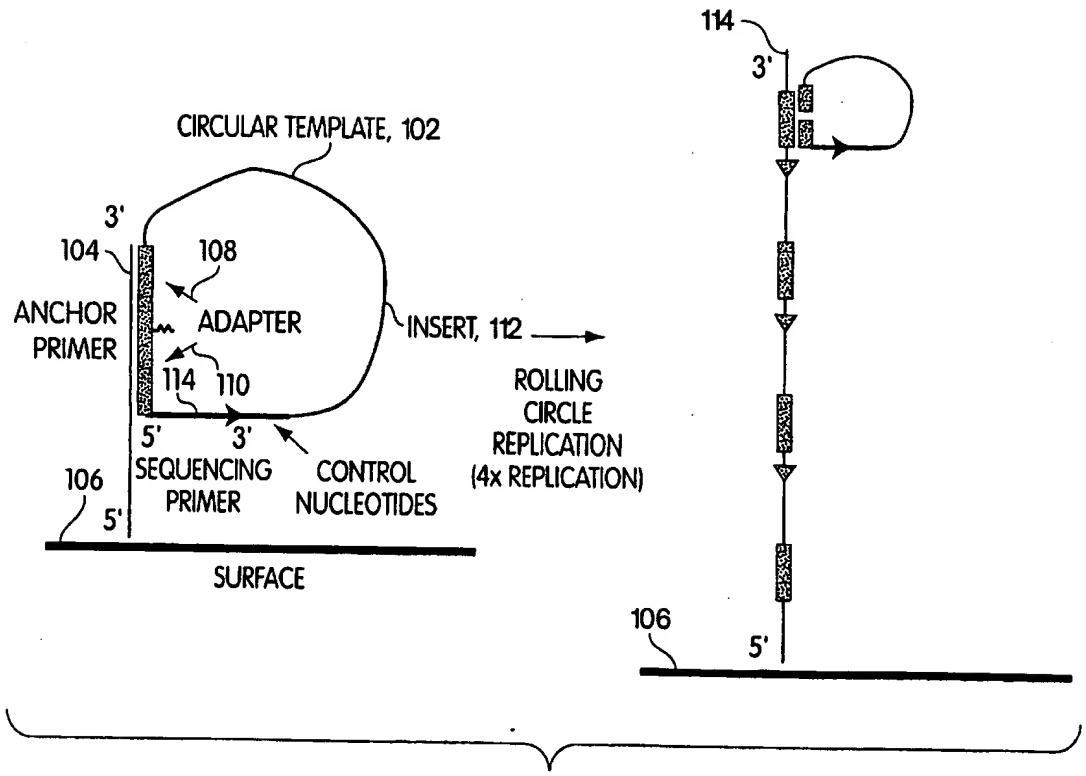

CIRCULAR TEMPLATE, 102
3'
104
108
ANCHOR PRIMER
ADAPTER
INSERT, 112
114
110
5'
3'
SEQUENCING PRIMER
CONTROL NUCLEOTIDES
106
5'
SURFACE
ROLLING CIRCLE REPLICATION (4x REPLICATION)
114
3'
106
5'

Fig. 1A

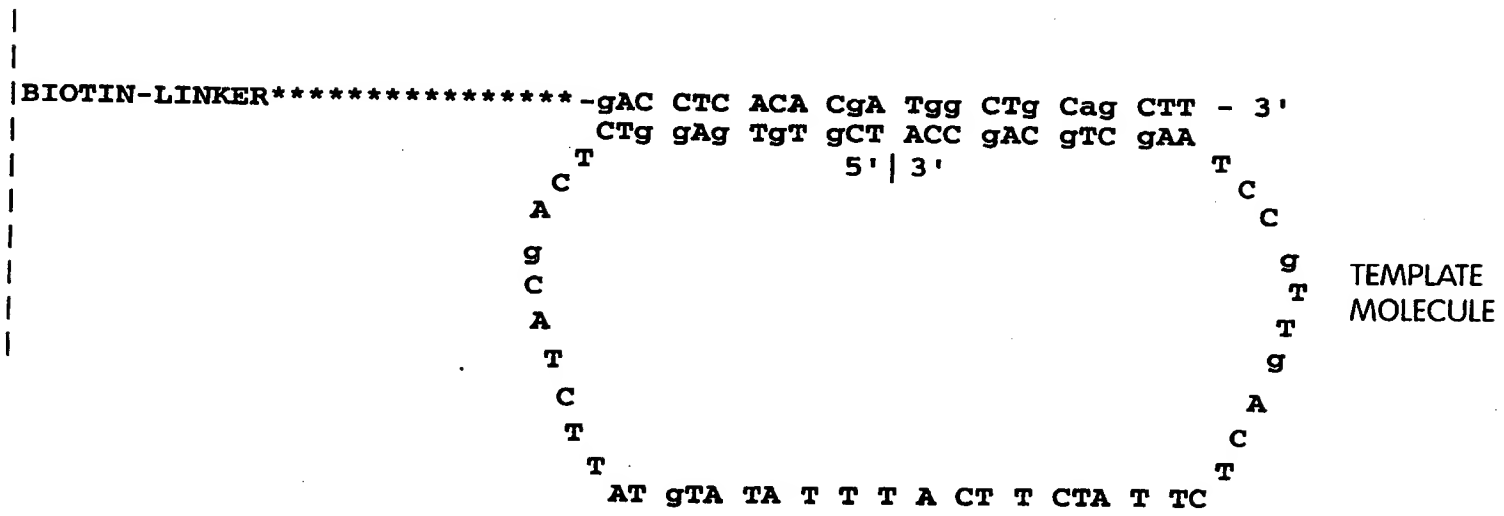

BIOTIN-LINKER**********-gAC CTC ACA CgA Tgg CTg Cag CTT - 3'
CTg gAg TgT gCT ACC gAC gTC gAA
5' | 3'
AT gTA TA T T T A CT T CTA T TC
TEMPLATE MOLECULE

Fig. 1B

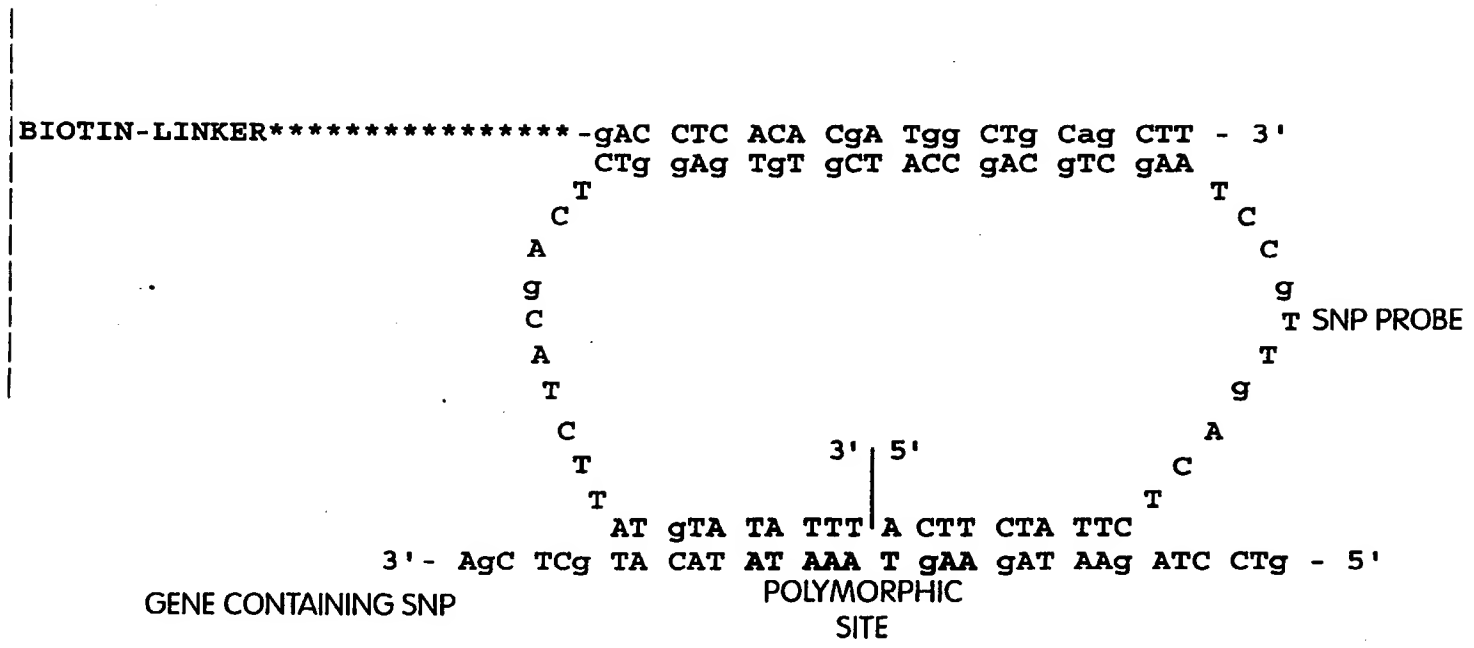

BIOTIN-LINKER**********-gAC CTC ACA CgA Tgg CTg Cag CTT - 3'
CTg gAg TgT gCT ACC gAC gTC gAA
T
C
A
g
C
A
T
C
T
T
T
C
C
g
T SNP PROBE
T
g
A
C
T
3' 5'
AT gTA TA TTT A CTT CTA TTC
3'- AgC TCg TA CAT AT AAA T gAA gAT AAg ATC CTg - 5'
GENE CONTAINING SNP
POLYMORPHIC SITE

Fig. 1C

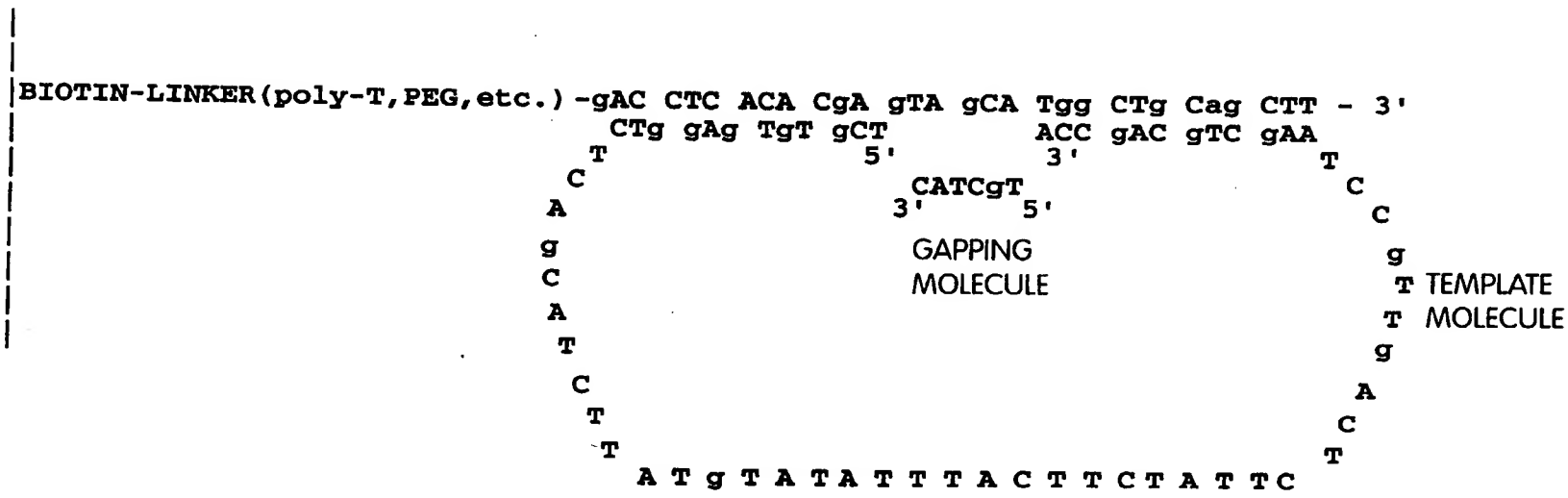

BIOTIN-LINKER(poly-T,PEG,etc.)-gAC CTC ACA CgA gTA gCA Tgg CTg Cag CTT - 3'
CTg gAg TgT gCT
ACC gAC gTC gAA
5'
3'
CATCgT
3'
5'
GAPPING
MOLECULE
TEMPLATE
MOLECULE
A T g T A T A T T T A C T T C T A T TC

Fig. 1D

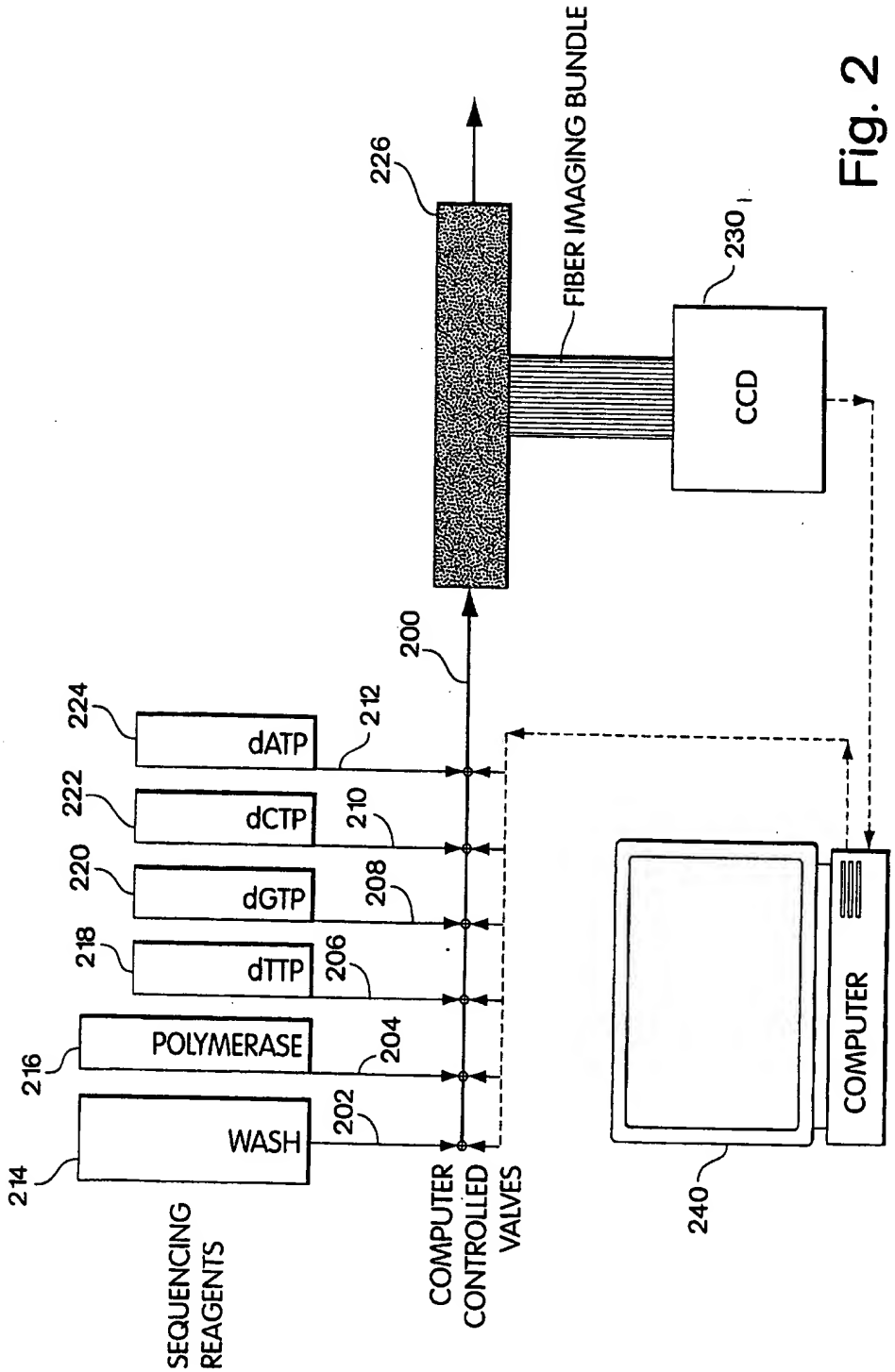

226
FIBER IMAGING BUNDLE
230
CCD
200
224
dATP
212
222
dCTP
210
220
dGTP
208
218
dTTP
206
216
POLYMERASE
204
214
WASH
202
SEQUENCING REAGENTS
COMPUTER CONTROLLED VALVES
COMPUTER
240

Fig. 2

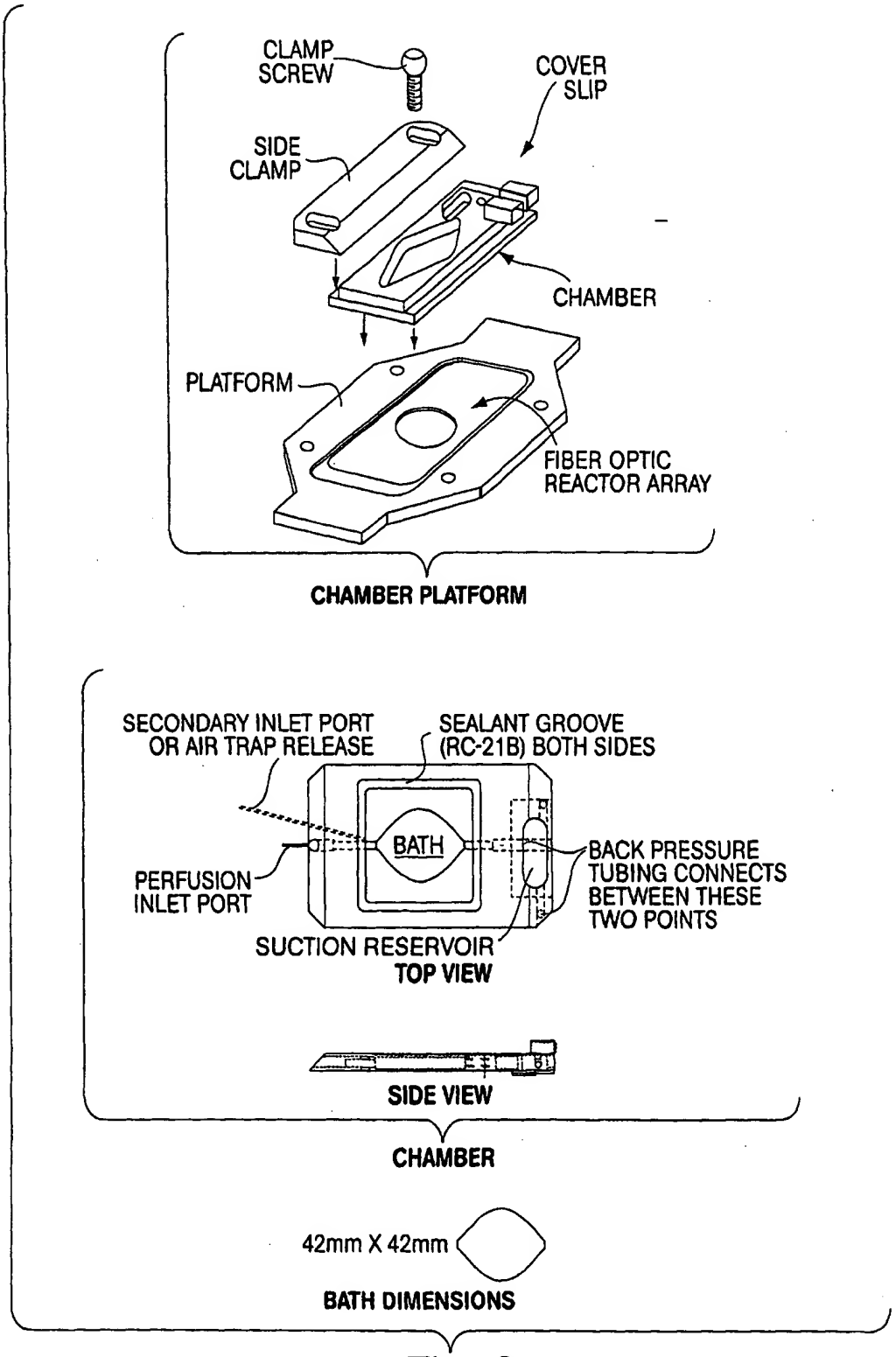

CLAMP SCREW
COVER SLIP
SIDE CLAMP
CHAMBER
PLATFORM
FIBER OPTIC REACTOR ARRAY
CHAMBER PLATFORM
SECONDARY INLET PORT OR AIR TRAP RELEASE
SEALANT GROOVE (RC-21B) BOTH SIDES
BATH
PERFUSION INLET PORT
BACK PRESSURE TUBING CONNECTS BETWEEN THESE TWO POINTS
SUCTION RESERVOIR
TOP VIEW
SIDE VIEW
CHAMBER
42mm X 42mm
BATH DIMENSIONS

Fig. 3

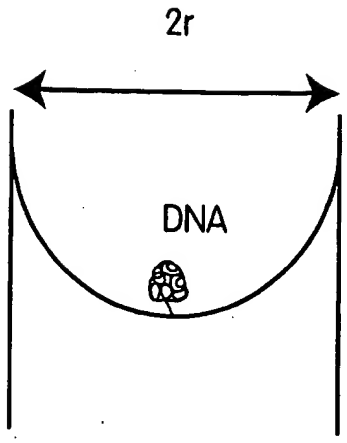


Fig. 4

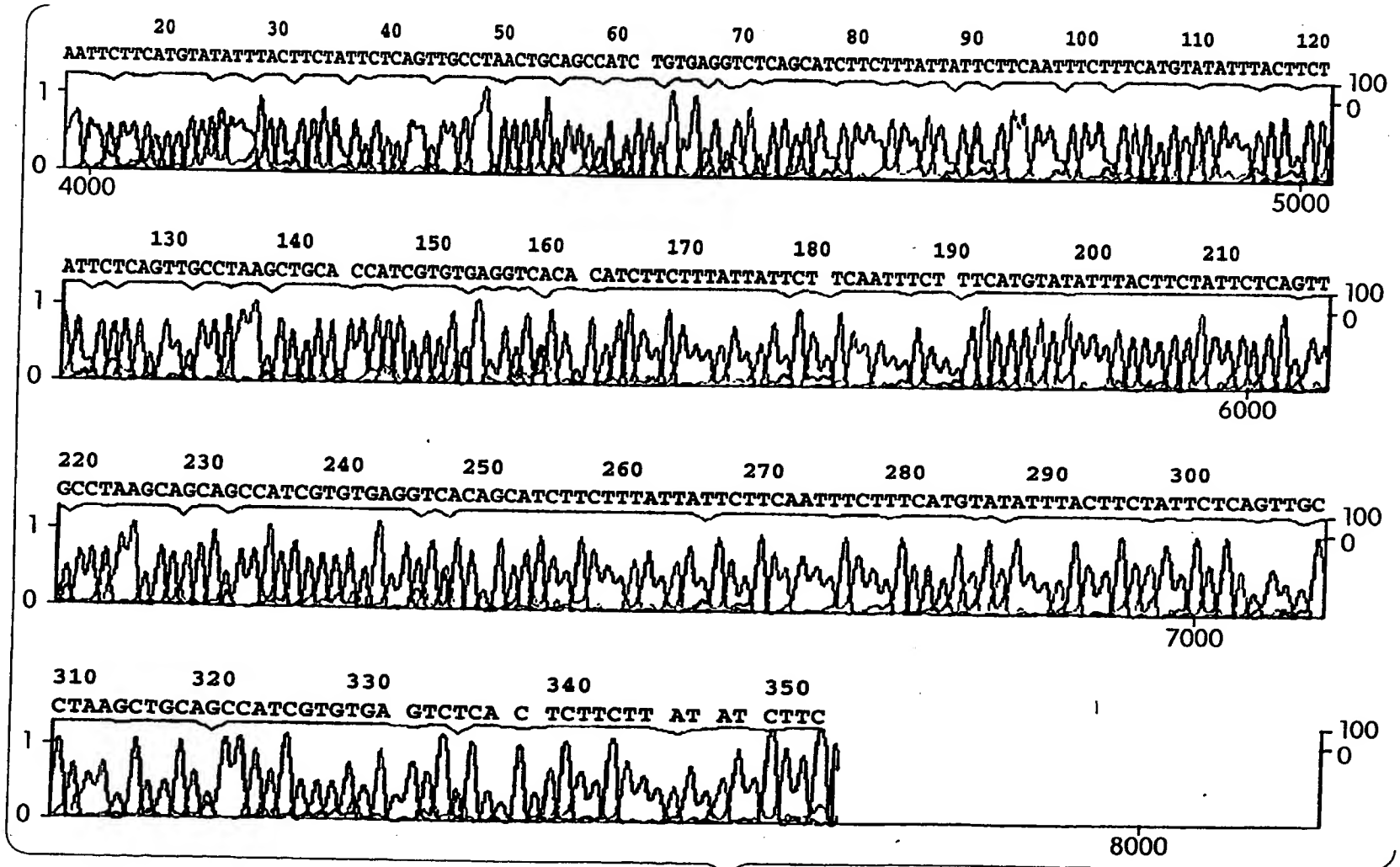

20 30 40 50 60 70 80 90 100 110 120
AATTCTTCATGTATATTTACTTCTATTCTCAGTTGCCTAACTGCAGCCATC TGTGAGGTCTCAGCATCTTCTTTATTATTCTTCAATTTCTTTCATGTATATTTACTTCT
1
0
4000
5000
100
0
130 140 150 160 170 180 190 200 210
ATTCTCAGTTGCCTAAGCTGCA CCATCGTGTGAGGTCACA CATCTTCTTTATTATTCT TCAATTTCT TTCATGTATATTTACTTCTATTCTCAGTT
1
0
6000
100
0
220 230 240 250 260 270 280 290 300
GCCTAAGCAGCAGCCATCGTGTGAGGTCACAGCATCTTCTTTATTATTCTTCAATTTCTTTCATGTATATTTACTTCTATTCTCAGTTGC
1
0
7000
100
0
310 320 330 340 350
CTAAGCTGCAGCCATCGTGTGA GTCTCA C TCTTCTT AT AT CTTC
1
0
8000
100
0

Fig. 5

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

35437 7590 02/07/2007

MINTZ LEVIN COHN FERRIS GLOVSKY & POPEO
666 THIRD AVENUE
NEW YORK, NY 10017

EXAMINER	
KIM, YOUNG J	
ART UNIT	PAPER NUMBER
1637	

DATE MAILED: 02/07/2007

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/814,338	03/21/2001	Jonathan M. Rothberg	21465-501 CIP2	6233

TITLE OF INVENTION: METHOD OF SEQUENCING A NUCLEIC ACID

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	YES	$700	$300	$0	$1000	05/07/2007

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the SMALL ENTITY status shown above.

If the SMALL ENTITY is shown as YES, verify your current SMALL ENTITY status:

A. If the status is the same, pay the TOTAL FEE(S) DUE shown above.

B. If the status above is to be removed, check box 5b on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and twice the amount of the ISSUE FEE shown above, or

If the SMALL ENTITY is shown as NO:

A. Pay TOTAL FEE(S) DUE shown above, or

B. If applicant claimed SMALL ENTITY status before, or is now claiming SMALL ENTITY status, check box 5a on Part B - Fee(s) Transmittal and pay the PUBLICATION FEE (if required) and 1/2 the ISSUE FEE shown above.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed and an extra copy of the form should be submitted. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Utility patents issuing on applications filed on or after Dec. 12, 1980 may require payment of maintenance fees. It is patentee's responsibility to ensure timely payment of maintenance fees when due.

 EXHIBIT B

Complete and send this form, together with applicable fee(s), to: Mail Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450
or Fax (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

35437 7590 02/07/2007

MINTZ LEVIN COHN FERRIS GLOVSKY & POPEO
666 THIRD AVENUE
NEW YORK, NY 10017

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being facsimile transmitted to the USPTO (571) 273-2885, on the date indicated below.

(Depositor's name)
(Signature)
(Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/814,338	03/21/2001	Jonathan M. Rothberg	21465-501 CIP2	6233

TITLE OF INVENTION: METHOD OF SEQUENCING A NUCLEIC ACID

APPLN. TYPE	SMALL ENTITY	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	YES	$700	$300	$0	$1000	05/07/2007

EXAMINER	ART UNIT	CLASS-SUBCLASS
KIM, YOUNG J	1637	435-006000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).
☐ Change of correspondence address (or Change of Correspondence Address form PTO/SB/122) attached.
☐ "Fee Address" indication (or "Fee Address" Indication form PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) the names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) the name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.
1 _____
2 _____
3 _____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document has been filed for recordation as set forth in 37 CFR 3.11. Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ☐ Individual ☐ Corporation or other private group entity ☐ Government

4a. The following fee(s) are submitted:
☐ Issue Fee
☐ Publication Fee (No small entity discount permitted)
☐ Advance Order - # of Copies _____

4b. Payment of Fee(s): **(Please first reapply any previously paid issue fee shown above)**
☐ A check is enclosed.
☐ Payment by credit card. Form PTO-2038 is attached.
☐ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment, to Deposit Account Number _____ (enclose an extra copy of this form).

5. **Change in Entity Status** (from status indicated above)
☐ a. Applicant claims SMALL ENTITY status. See 37 CFR 1.27.
☐ b. Applicant is no longer claiming SMALL ENTITY status. See 37 CFR 1.27(g)(2).

NOTE: The Issue Fee and Publication Fee (if required) will not be accepted from anyone other than the applicant; a registered attorney or agent; or the assignee or other party in interest as shown by the records of the United States Patent and Trademark Office.

Authorized Signature _____ Date _____

Typed or printed name _____ Registration No. _____

This collection of information is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450.

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/814,338	03/21/2001	Jonathan M. Rothberg	21465-501 CIP2	6233

35437 7590 02/07/2007

MINTZ LEVIN COHN FERRIS GLOVSKY & POPEO
666 THIRD AVENUE
NEW YORK, NY 10017

EXAMINER	
KIM, YOUNG J	
ART UNIT	PAPER NUMBER
1637	

DATE MAILED: 02/07/2007

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(application filed on or after May 29, 2000)

The Patent Term Adjustment to date is 0 day(s). If the issue fee is paid on the date that is three months after the mailing date of this notice and the patent issues on the Tuesday before the date that is 28 weeks (six and a half months) after the mailing date of this notice, the Patent Term Adjustment will be 0 day(s).

If a Continued Prosecution Application (CPA) was filed in the above-identified application, the filing date that determines Patent Term Adjustment is the filing date of the most recent CPA.

Applicant will be able to obtain more detailed information by accessing the Patent Application Information Retrieval (PAIR) WEB site (http://pair.uspto.gov).

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

PTOL-85 (Rev. 07/06) Approved for use through 04/30/2007.

Notice of Allowability	Application No. 09/814,338	Applicant(s) ROTHBERG ET AL.	
	Examiner Young J. Kim	Art Unit 1637	

-- ***The MAILING DATE of this communication appears on the cover sheet with the correspondence address***--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☒ This communication is responsive to *the RCE received on December 28, 2006*.

2. ☒ The allowed claim(s) is/are *56-61,64-68,84-93 and 96-99*.

3. ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   a) ☐ All b) ☐ Some* c) ☐ None of the:
      1. ☐ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
      3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

   * Certified copies not received: _____.

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

4. ☐ A SUBSTITUTE OATH OR DECLARATION must be submitted. Note the attached EXAMINER'S AMENDMENT or NOTICE OF INFORMAL PATENT APPLICATION (PTO-152) which gives reason(s) why the oath or declaration is deficient.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   (a) ☐ including changes required by the Notice of Draftsperson's Patent Drawing Review (PTO-948) attached
      1) ☐ hereto or 2) ☐ to Paper No./Mail Date _____.
   (b) ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____.

   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)

1. ☐ Notice of References Cited (PTO-892)
2. ☐ Notice of Draftperson's Patent Drawing Review (PTO-948)
3. ☒ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date 12/28/06
4. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material
5. ☐ Notice of Informal Patent Application
6. ☐ Interview Summary (PTO-413), Paper No./Mail Date _____ .
7. ☐ Examiner's Amendment/Comment
8. ☒ Examiner's Statement of Reasons for Allowance
9. ☐ Other _____.

Young J. Kim
Primary Examiner
Art Unit: 1637
2-5-07

DETAILED ACTION

Examiner Reasons for Allowance

The diameter range, the depth, and well-depth of the claimed cavitated fiber optic wafer, said wafer being formed from a fused bundle of a plurality of individual optical fibers, are deemed non-obvious over the general teaching provided for by Chee et al. (of record), based on Margulies Declaration, as each of the above-mentioned parameters are critical to the laminar flow of the reaction reagents, allowing for rapid diffusion and washing out of the nucleotides, enabling a high cycle sequencing (see page 3 bottom paragraph to page 4, top paragraph; Margulies Declaration).

Hence, the parameters are no longer deemed arbitrary, not involving routine optimization, as the parameters are specifically adapted to the way in which the disclosed invention operates.

Inquiries

Any inquiry concerning this communication or earlier communications from the Examiner should be directed to Young J. Kim whose telephone number is (571) 272-0785. The Examiner is on flex-time schedule and can best be reached from 8:30 a.m. to 4:30 p.m (M-W and F). The Examiner can also be reached via e-mail to Young.Kim@uspto.gov. However, the office cannot guarantee security through the e-mail system nor should official papers be transmitted through this route.

If attempts to reach the Examiner by telephone are unsuccessful, the Examiner's supervisor, Dr. Gary Benzion, can be reached at (571) 272-0782.

Papers related to this application may be submitted to Art Unit 1637 by facsimile transmission. The faxing of such papers must conform with the notice published in the Official Gazette, 1156 OG 61 (November 16, 1993) and 1157 OG 94 (December 28, 1993) (see 37 CFR 1.6(d)). NOTE: If applicant does submit a paper by FAX, the original copy should be retained by applicant or applicant's representative. NO DUPLICATE COPIES

SHOULD BE SUBMITTED, so as to avoid the processing of duplicate papers in the Office. All official documents must be sent to the Official Tech Center Fax number: (571) 273-8300. For Unofficial documents, faxes can be sent directly to the Examiner at (571) 273-0785. Any inquiry of a general nature or relating to the status of this application should be directed to the Group receptionist whose telephone number is (571) 272-1600.

Young J. Kim
Primary Examiner
Art Unit 1637
2/5/2007

YOUNG J. KIM
PRIMARY EXAMINER

YJK

Please type a plus sign (+) in this box [+]

PTO/SB (12-97)
OMB 0651-0031
Patent and Trademark Office: U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Modified Form 1449/PTO **SUPPLEMENTAL INFORMATION DISCLOSURE STATEMENT BY APPLICANT** (use as many sheets as necessary)		
	Application Number	09/814,338
	Filing Date	03/21/01
	First Named Inventor	Rothberg
	Group Art Unit / Confirmation Number	1637 / 6233
	Examiner Name	Young J. Kim
	Attorney Docket Number	21465-501 CIP2

U.S. PATENT DOCUMENTS							
Exam Initials	Cite No.	U.S. Patent Document No.	Issue Date	Name of Patentee(s) or Applicant(s)	Class	Sub Class	Filing Date If Appropriate
YK	A184	4,402,568	09/06/83	Kulich et al.	350	96.16	02/09/81

FOREIGN PATENT DOCUMENTS							
Exam Initials	Cite No.	Foreign Patent Document Office	Number	Name of Patentee(s) or Applicant(s)	Date of Publication	English Yes	No

OTHER PRIOR ART – NON PATENT LITERATURE DOCUMENTS		
Exam Initials	Cite No.	Name of Author, Title (when appropriate), Publication, Volume, Page(s), Date, Etc.

*a copy of this reference is not provided as it was previously cited by or submitted to the office in a prior application, Serial No. _____, filed _____, and relied upon for an earlier filing date under 35 U.S.C. §120 (continuation, continuation-in-part, and divisional applications).

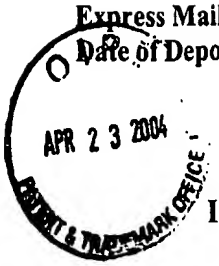
Examiner Signature	/Young J. Kim/	Date Considered	01/31/2007

EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609; Draw line through citation if not in conformance and not considered.

Include copy of this form with next communication to applicant.

Express Mail Label No.: EV287336791US
Date of Deposit: April 23, 2004

Attorney Docket No. 21465-501 CIP2



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS:	Jonathan M. Rothberg, et al.		
ASSIGNEE:	CuraGen Corporation		
SERIAL NUMBER:	09/814,338	EXAMINER:	Young J. Kim
FILING DATE:	March 21, 2001	ART UNIT:	1637
FOR:	METHOD OF SEQUENCING A NUCLEIC ACID		

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

AMENDMENT AND RESPONSE TO NON-FINAL OFFICE ACTION DATED NOVEMBER 6, 2003

In response to the Non-Final Office Action mailed November 6, 2003 in the above-identified case, Applicants submit this Amendment and Response. Also submitted is a Request for a Three-Month Extension of Time and the fee required under 37 C.F.R. § 1.17(a)(3). Applicants further submit a supplemental Information Disclosure Statement, Form PTO-1449, and the fee required under 37 C.F.R. § 1.17(p). Applicants respectfully request entry and consideration of this Amendment and accompanying papers in the application.

Amendment

Please amend the application as follows:

Amendments to the Specification begin on page 2 of this paper.

Amendments to the Claims are reflected in the listing of claims, which begins on page 3 of this paper.

Remarks/Arguments begin on page 7 of this paper.

EXHIBIT C

This listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

1.-55. (Cancelled)

56. (Currently amended) A substrate for analyzing a nucleic acid, the substrate comprising:

a cavitated fiber optic wafer formed from a fused bundle of a plurality of individual optical fibers, each individual optical fiber having a diameter between 3 and 100 μm, the wafer comprising a top surface and a bottom surface, the top surface comprising at least 10,000 wells, wherein said wells are etched into the top surface of the cavitated fiber optic wafer and wherein the thickness of the wafer between the top surface and the bottom surface is between 0.5 mm and 5.0 mm in thickness; wherein the depth of each well ranges from between one half the diameter of an individual optical fiber and three times the diameter of an individual optical fiber; and wherein a plurality of wells on the top surface of the cavitated wafer have a nucleic acid ~~disposed~~ therein; and

a plurality of beads ~~disposed~~ within wells on the top surface of the cavitated wafer, said beads having a pyrophosphate sequencing reagent attached thereto.

57. (Previously amended) The substrate of claim 56, wherein the nucleic acid is immobilized on the wells or on said beads.

58. (Previously amended) The substrate of claim 56, wherein the diameter of each individual optical fiber in the cavitated wafer is between 6-50 μm.

59. (Previously amended) The substrate of claim 58, wherein the fiber optic surface includes two or more nucleic acids separated by approximately 10 μm to approximately 200 μm.

60. (Previously amended) The substrate of claim 58, wherein the fiber optic surface includes two or more nucleic acids separated by approximately 10 μm to approximately 150 μm.

Amendments to the Specification:

Please replace the paragraph on page 43, lines 16-21 with the following paragraph:

In addition, the diffusion constants DP for PPi and DA for ATP must also be specified. These values may be estimated from the following exemplar diffusion constants for biomolecules in a dilute water solution (see Weisiger, 1997. Impact of Extracellular and Intracellular Diffusion on Hepatic Uptake Kinetics Department of Medicine and the Liver Center, University of California, San Francisco, California, USA, ~~dickw@itsa.ucsf.edu~~, ~~http://dickw.ucsf.edu/papers/goresky97/chapter.html~~ available online at hypertext transfer protocol://dickw.ucsf.edu/papers/goresky97/chapter.html).

Please insert the following paragraph after the paragraph ending on line 17 on page 33:

-- In one aspect, the invention embodies an apparatus for processing a plurality of analytes, the apparatus comprising: a flow chamber having therein a substrate comprising a plurality of cavitated surfaces that have thereon nucleic acid molecules; fluid means for delivering processing reagents from one or more reservoirs to the flow chamber so that the analytes anchored to the plurality of microparticles are exposed to the reagents; and detection means for detecting a sequence of optical signals from each microparticle of the plurality, each optical signal of the sequence being indicative of an interaction between a processing reagent and the analyte anchored thereto, wherein the detection means is in communication with the cavitated surfaces. The detection means may further comprise signal tracking means for correlating said optical signals from each of the microparticles in each of the digital images to form a sequence. The signal tracking means may comprise a CCD camera, and the analyte may comprise DNA. --

61. (Previously amended) The substrate of claim 58, wherein the fiber optic surface includes two or more nucleic acids separated by approximately 150 μm.

62.-63. (Cancelled)

64. (Previously amended) The substrate of claim 56 wherein the wafer further comprises 10^3 or more groups of nucleic acid sequences in said wells.

65. (Previously amended) The substrate of claim 64, wherein said substrate comprises 10^4 or more different groups of nucleic acid sequences in discrete known regions.

66. (Previously amended) The substrate of claim 64, wherein said substrate comprises 10^5 or more different groups of nucleic acid sequences in discrete known regions.

67. (Previously amended) The substrate of claim 64, wherein the nucleic acid sequences are attached to the wells or beads by a linker.

68. (Previously amended) The substrate of claim 64, wherein the nucleic acid sequences are covalently attached to the wells or beads.

69.- 83. (Cancelled)

84. (Currently amended) An apparatus for processing a plurality of nucleic acids, the apparatus comprising:

a flow chamber having disposed therein a cavitated fiber optic wafer;

a cavitated fiber optic wafer formed from a fused bundle of a plurality of individual optical fibers, each individual optical fiber having a diameter between 3 and 100 μm, the wafer comprising a top surface and a bottom surface, the top surface comprising at least 10,000 wells, wherein said wells are etched into the top surface of the cavitated fiber optic wafer and wherein the thickness of the wafer between the top surface and the bottom surface is between 0.5 mm and

5.0 mm in thickness; wherein the depth of each well ranges from between one half the diameter of an individual optical fiber and three times the diameter of an individual optical fiber; and wherein a plurality of wells on the on the top surface of the cavitated wafer have a nucleic acid ~~disposed~~ therein;

a plurality of beads ~~disposed~~ within wells on the top surface of the cavitated wafer, said beads having a pyrophosphate sequencing reagent attached thereto;

fluid means for delivering additional pyrophosphate sequencing reagents, including sequential delivery of nucleotide triphosphates, from one or more reservoirs to the flow chamber so ~~the~~ nucleic acids ~~disposed on beads~~ in the wells on the top surface of the fiber optic wafer are exposed to the reagents; and

detection means for detecting optical signals from each well, wherein said detection means is in communication with the wells, each optical signal being indicative of reaction of the pyrophosphate sequencing reagents with the nucleic acid in a well.

85. (Previously amended) The apparatus of claim 84, wherein the diameter of each individual optical fiber in the cavitated wafer is between 6-50 μm.

86. (Previously amended) The apparatus of claim 85, wherein said detection means is a CCD camera.

87. (Previously amended) The apparatus of claim 84, wherein the nucleic acid is DNA.

88. (Previously amended) The substrate of claim 56 wherein the substrate has a polished fiber optic surface opposite to the cavitated fiber optic surface.

89. (Previously added) The substrate of claim 88 wherein the polished surface allows for optical coupling to a second optical fiber.

90. (Previously amended) The substrate of claim 56 wherein the cavitated fiber optic wafer is coated.

91. (Previously added) The substrate of claim 90 wherein the coating is selected from the group consisting of plastic, gold layers, organosilane reagents, photoreactive linkers, hydrophilic polymer gels and pluronic polymers.

92. (Currently amended) The substrate of claim 56 wherein said pyrophosphate sequencing reagent is luciferase.

93. (Currently amended) The substrate of claim 56 wherein said pyrophosphate sequencing reagent is sulfurylase.

94.-95. (Cancelled)

96. (Previously amended) The apparatus of claim 84 wherein the cavitated fiber optic wafer is coated.

97. (Previously amended) The apparatus of claim 96 wherein the coating is selected from the group consisting of plastic, gold layers, organosilane reagents, photoreactive linkers, hydrophilic polymer gels and pluronic polymers.

98. (Previously amended) The apparatus of claim 84 wherein said pyrophosphate sequencing reagent is luciferase.

99. (Previously amended) The apparatus of claim 84 wherein said pyrophosphate sequencing reagent is sulfurylase.

100. (New) The apparatus of claim 84, wherein the nucleic acid is immobilized on the wells or on said beads.

REMARKS

Amendments

Claims 56-61, 64-68, 84-93, and 96-100 are currently pending in the application.

Claims 1-55 have been cancelled without prejudice or disclaimer solely to comply with the requirement to cancel the non-elected claims in the application. Applicants reserve the right to present the cancelled claims in a co-pending application.

Claims 56, 84, 92, 93, and 98 have been amended for clarity.

Specifically, claim 56 has been amended:

- to replace recitation of "nucleic acid disposed therein" with "nucleic acid therein;" and
- to replace the recitation "plurality of beads disposed within wells" with "plurality of beads within wells" (see, *inter alia*, page 31, lines 21-29).

Claim 84 has been amended:

- to replace recitation of "a nucleic acid disposed therein" with "a nucleic acid therein;"
- to replace recitation of "plurality of beads disposed within wells" with "plurality of beads within wells;"
- to replace recitation of "the nucleic acids disposed on beads in the wells" with "nucleic acids in the wells;" and
- to replace recitation of "delivering pyrophosphate sequencing reagents" with "delivering additional pyrophosphate sequencing reagents" (see, *inter alia*, page 31, lines 21-29).

Claims 92, 93, and 98 have been amended to replace recitation of "said sequencing reagent" with "said pyrophosphate sequencing reagent" (see, *inter alia*, page 31, lines 21-29).

Claim 100 has been added to more fully encompass Applicants' invention. Claim 100 includes recitation of "nucleic acid is immobilized on the wells or on said beads" (see, *inter alia*, page 31, lines 21-29 and original claims 84-87 (discussed below)).

The paragraph on page 43, lines 16-21 has been amended to replace recitation of "http" with "hypertext transfer protocol," which is the corresponding phrase for this acronym.

A new paragraph has been inserted on page 33, line 17 to include the subject matter disclosed in original claims 84-87 (discussed below).

These amendments are supported by the application as originally filed, and do not constitute new matter. Specific support for the amendments is shown in parentheses, above. Entry of these amendments is respectfully requested.

Priority

The Examiner has stated that the claimed subject matter lacks support under 35 U.S.C. §112, first paragraph (Office Action, page 2). The Examiner has thereby concluded that the effective priority date for the instant application is the filing date of March 21, 2001 (Office Action, page 2). Applicants respectfully assert that the claims of the instant application are supported by the disclosure of priority document U.S. Application Serial No. 09/664,197, filed September 18, 2000. As illustration, Applicants point to the specific disclosure of Application Serial No. 09/664,197, as follows.

- Page 4, line 4 to page 7, line 23, discloses, *inter alia*, the cavitated fiber optic surface, the bundled fiber optic surfaces, apparatuses for sequence analysis, pyrophosphate sequencing reagents, linkage of nucleic acids to the support surface, the distance of the nucleic acids on the claimed substrate, the number of wells, and the sequencing of PCR products (discussed below).
- Figures 1-4 disclose, *inter alia*, the CCD and computer, the fiber optic substrate, and the wells of the fiber optic substrate;
- Page 8, line 5 to page 9, line 18, discloses, *inter alia*, linkage of nucleic acids to the substrate surface and the sequencing of PCR products (discussed below);
- Page 10, line 21 to page 15, line 6, discloses, *inter alia*, linkage of nucleic acids to the substrate surface;
- Page 24, line 9 to page 32, line 24, discloses, *inter alia*, pyrophosphate sequencing reagents (e.g., sulfurylase and luciferase), sequential addition of dNTPs, and the immobilization of sequencing reagents on beads;
- Page 32, line 26 to page 38, line 11, discloses, *inter alia*, the diameter of the individual fibers, the thickness of the wafer, the depth of the wells, the optional connection to a second fiber optic bundle, and the coating of the wafer;
- Examples 1 and 2 disclose, *inter alia*, the tandem amplification and sequencing of a nucleic acid on the same wafer surface;

- Original claims 59-62 include the same disclosure as original claims 84-87 of the instant application (discussed below).

Applicants note that the Examiner has not identified any specific subject matter of the claimed invention that allegedly lacks support in priority document U.S. Application Serial No. 09/664,197. In addition, the Examiner has not specified the individual claims that are alleged to lack such support. As listed above, Applicants have pointed to exemplary disclosure in priority document U.S. Application Serial No. 09/664,197 which provides support for the subject matter of the claims of instant application under 35 U.S.C. §112, first paragraph. It is respectfully asserted that the pending claims of the instant application are entitled to a priority date of September 18, 2000.

Election/Restrictions

The Examiner has required Applicants to cancel claims 1-55 of the instant application, as such claims are directed to non-elected subject matter (Office Action, page 2). Applicants have canceled claims 1-55 as a result of this Amendment, but reserve the right to present the cancelled claims in a co-pending application (see above).

Information Disclosure Statements

The Examiner has provided Applicants with a signed copy of the Modified Form PTO-1449 filed on February 1, 2002, as well as newly signed copies of the Modified Form PTO-1449 filed on June 12, 2003 and the Modified Form PTO-1449 filed on September 5, 2003 indicating consideration of all of the listed publications (Office Action, pages 2-3). Applicants acknowledge receipt of these papers in connection with the instant application.

Specification

The Examiner has objected to the specification for containing a hyperlink or other form of browser-executable code (Office Action, page 3). In particular, the Examiner pointed to the hyperlink on page 43, lines 20-21 of the instant application (Office Action, page 3). Applicants have amended the specification to deactivate the hyperlink (see above). Specifically, the acronym "http" has been replaced with the corresponding phrase "hypertext transfer protocol." This amendment is supported by the original application, as filed, and constitutes no new matter

(see above). It is believed that deactivation of the hyperlink in the specification obviates this ground of objection, and reconsideration is respectfully requested.

Rejection under 35 U.S.C. §112, First Paragraph

The Examiner has rejected claims 56-61, 64-68, 84-93, and 96-99 under 35 U.S.C. §112, first paragraph as allegedly failing to comply with the written description requirement (Office Action, page 4). The Examiner states that the rejected claims contain new matter, i.e., subject matter which the Examiner has deemed not described in the specification in such a way to reasonably convey to one skilled in the art that the inventors had possession of the claimed invention at the time the application was filed (Office Action, page 4). In particular, it is alleged that there is no support in the specification for the subject matter of 1) sequencing reagents immobilized on beads (Office Action, page 5); or 2) nucleic acids attached to beads (Office Action, page 6). Applicants respectfully traverse this rejection as follows.

1. **Support for Pyrophosphate Sequencing Reagents Attached to Beads**

Applicants respectfully assert that the instant application fully supports the claimed subject matter directed to sequencing reagents immobilized on beads. As presently amended, claims 56 and 84 recite the phrase "beads having a pyrophosphate sequencing reagent attached thereto." In support of this recitation, Applicants have pointed to page 31, lines 21-29 of the originally filed specification, which states:

> In various embodiments, some components of the reaction are immobilized, while other components are provided in solution. For example, in some embodiments, the enzymes utilized in the pyrophosphate sequencing reaction (e.g., sulfurylase, luciferase) may be immobilized if desired onto the solid support. Similarly, one or more or of the enzymes utilized in the pyrophosphate sequencing reaction, e.g., sulfurylase, luciferase may be immobilized at the termini of a fiber optic reactor array. Other components of the reaction, e.g., a polymerase (such as Klenow fragment), nucleic acid template, and nucleotides can be added by flowing, spraying, or rolling. In still further embodiments, one more of the reagents used in the sequencing reactions is *delivered on beads* (Emphasis added).

When read in context, this paragraph describes various ways that the components of the sequencing reaction can be presented. According to this paragraph, a reaction component 1) can

be immobilized or provided in solution; 2) (an enzyme) can be immobilized on the solid support; 3) (an enzyme) can be immobilized at the array termini; 4) can be added by flowing, spraying, or rolling; or 5) can be delivered on beads.

In the Office Action, the Examiner states that use of the language "*delivered on* beads" indicates that the reagents are specifically *not* immobilized on the beads (Office Action, page 5). Yet, the phrase "delivered on" is commonly used in scientific publications to refer to delivery via an attachment or linkage. Applicants point to Exhibits 1-7, attached hereto. In particular:

- Rhoades et al., 2003, *Mol. Microbiol.* 48:875-888 (Exhibit 1) explicitly refers to BCG lipids "delivered on" polystyrene microspheres (see, *inter alia*, summary; page 881, left column; and page 884, left column);

- Sedegah et al., 2000, *J. Immunol.* 164:5905-5912 (Exhibit 2) explicitly refers to DNA "delivered on" gold particles (see, *inter alia*, page 5911, left column);

- Leister et al., 1996, *Proc. Natl. Acad. Sci. USA* 93:15497-15502 (Exhibit 3) explicitly refers to a LUC reporter gene "delivered on" biolistic particles (see, *inter alia*, page 15501, right column);

- Bowes et al., 2002, *Infect. Immun.* 70:5008-5018 (Exhibit 4) explicitly refers to a GD3 gangloside "delivered on" lipopolysaccharides (see, *inter alia*, page 5013, bottom right column to page 5014, top left column);

- Reeves et al., 2002, *J. Biol. Chem.* 277:9155-9159 (Exhibit 5) explicitly refers to a replacement cassette "delivered on" a KC515 phage vector (see, *inter alia*, page 9157, legend for Figure 2);

- Legge et al., 1997, *J. Exp. Med.* 185:1043-1053 (Exhibit 6) explicitly refers to a PLP-LR peptide "delivered on" an Ig chimera (see, *inter alia*, page 1050, bottom left column to top right column); and

- Burton and Gray, 1995, *Br. J. Cancer* 71:322-325 (Exhibit 7) explicitly refers to radiation doses "delivered on" microspheres (see, *inter alia*, summary and page 323, top left column).

In each of these articles, the authors have employed the phrase "delivered on" to mean "delivery of '*X*' via attachment or linkage to '*Y*'." In view of this common usage, it is clear that "one or more of the reagents...is *delivered on beads*," means "one or more of the reagents...is delivered *via attachment or linkage to beads*" as disclosed in the originally filed specification.

The Examiner argues that Applicants do not repeat the disclosure of page 31 ("one more of the reagents used in the sequencing reactions is delivered on beads") elsewhere in the

specification. However, Applicants are not required to repeat a disclosure throughout the instant application. Further, Applicants respectfully submit that the Examiner's interpretation of the specification would lead to an illogical result.

If the Examiner's interpretation were accepted, then *none* of the key components of the reaction could be deemed attached to beads, since the Examiner argues that the specification fails to support *either* sequencing reagents immobilized on beads (Office Action, page 5) *or* nucleic acids attached to beads (Office Action, page 6). This is clearly an erroneous outcome given the disclosure on page 31, lines 21-29 of the instant application (excerpted above). Applicants respectfully submit that the originally filed specification fully supports claims 56 and 84, which recite "beads having a pyrophosphate sequencing reagent attached thereto."

2. **Support for Nucleic Acids Attached to Beads**

Applicants respectfully assert that the instant application fully supports the claimed subject matter directed to nucleic acids attached to beads. As presented herein:

- Claim 57 recites "nucleic acid is immobilized on the wells or on said beads;"
- Claim 67 recites "nucleic acid sequences are attached to the wells or beads;"
- Claim 68 recites "nucleic acid sequences are...attached to the wells or beads;" and
- New claim 100 recites "nucleic acid is immobilized on the wells or on said beads."

For clarity, independent claim 84 has been amended to delete the recitation "the nucleic acids disposed on beads" (see above).

The Examiner states that the application teaches *only* that biotin-avidin binding can be used to immobilize nucleic acids to the reaction wells of the apparatus (Office Action, pages 5-6). Yet, Applicants point to page 39, lines 21-22 of the originally filed specification, which indicates that previous experiments had shown that PCR products were capable of being conjugated to streptavidin-coated magnetic beads.

In addition, the Examiner states that disclosure of "nucleic acids attached to beads" is nowhere to be found in the application (Office Action, page 5). Yet, Applicants point to page 31, lines 21-29 of the originally filed specification (excerpted above), and to original claims 84-87, which state:

84. An apparatus for processing a plurality of analytes, the apparatus comprising:

a flow chamber having disposed therein a substrate comprising a plurality of cavitated surfaces, said cavitated surfaces having disposed thereon nucleic acid molecules;

fluid means for delivering processing reagents from one or more reservoirs to the flow chamber so that the analytes anchored to the plurality of microparticles are exposed to the reagents; and

detection means for detecting a sequence of optical signals from each microparticle of the plurality, each optical signal of the sequence being indicative of an interaction between a processing reagent and the analyte anchored thereto, wherein said detection means is in communication with the cavitated surfaces.

85. The apparatus of claim 85, wherein said detection means further comprises signal tracking means for correlating said optical signals from each of said microparticles in each of said digital images to form for each said microparticle of said plurality a sequence of said optical signals.

86. The apparatus of claim 87, wherein said signal tracking means is a CCD camera.

87. The apparatus of claim 86, wherein said analyte is DNA. (Emphasis added).

These original claims describe various features of the sequencing apparatus, including cavitated surfaces (wells) with nucleic acids (e.g., DNA) anchored to microparticles (beads). As a result of this amendment, Applicants have incorporated original claims 84-87 in paragraph form in the specification (see above). The added paragraph does not constitute new matter, since the claims as filed in the original specification are part of the disclosure (MPEP §2163.06(III)). Moreover, the specification may be amended to include the subject matter of an originally filed claim (MPEP §2163.06(III)).

In addition, Applicants note that mere rephrasing of a passage does not constitute new matter (MPEP §2163.07(I)). As seen in scientific publications, the terms "microparticles" and "beads" are used interchangeably. Applicants point to Exhibits 8-13, attached hereto:

- Schweitzer et al., 2000, *Proc. Natl. Acad. Sci. USA* 97:10113-10119 (Exhibit 8; see, *inter alia*, page 10114, right column and page 10116, right column);
- Spiro et al., 2002, *Appl. Environ. Microbiol.* 68:1010-1013 (Exhibit 9; see, *inter alia*, page 1010, right column and page 1012, right column);

- Borowitz et al., 1997, *Blood* 89:3960-3966 (Exhibit 10; see, *inter alia*, page 3960, right column);
- Simone et al., 1999, *Am. J. Pathol.* 156:445-452 (Exhibit 11; see, *inter alia*, page 446, right column);
- Chen et al., 1999, *Clin. Chem.* 45:1693-1694 (Exhibit 12; see, *inter alia*, page 1693, right column and page 1694, right column); and
- de Bruin et al., 2000, *J. Gen. Virol.* 81:1529-1537 (Exhibit 13; see, *inter alia*, page 1532, left column).

Applicants respectfully submit that the originally filed application supports claims 57, 67, 68, and new claim 100, which recite "nucleic acid...immobilized on the wells or on said beads" or "nucleic acid sequences...attached to the wells or beads."

In view of all of the foregoing, Applicants conclude that independent claims 56 and 84 and corresponding dependent claims 57-61, 64-68, 84-87, 88-93, 96-100 do not contain new matter. Withdrawal of this rejection is respectfully requested.

3. **Telephone Interview Regarding Rejection under 35 U.S.C. §112, First Paragraph**

On November 19, 2003, Applicants' undersigned attorney and agent initiated a telephone interview with the Examiner to ask for clarification regarding the rejection set forth under 35 U.S.C. §112, first paragraph. The Examiner stated that it was unclear as to whether claims 56 and 84 included limitations for beads that were simultaneously attached to substrate and pyrophosphate sequencing reagents. The Examiner asked that the claims be amended to clarify this issue. Applicants' undersigned attorney and agent agreed to make amendments to claims 56 and 84 for clarity. As indicated above, claims 56 and 84 have been amended to change the recitation of "beads disposed within wells" to "beads within wells." While the term "disposed" is not limited by its usage in the instant application to mean "attached," Applicants wish to provide clarification as required by the Examiner. Withdrawal of the new matter rejection is therefore respectfully requested.

Rejection under 35 U.S.C. §112, Second Paragraph

The Examiner has rejected claims 84, 85-87, 92, 93, 96-99 under 35 U.S.C. §112, second paragraph, as allegedly indefinite for failing to particularly point out and distinctly claim the

subject matter Applicants' regard as the invention (Office Action, page 6). The Examiner has stated:

- In claim 84, the phrase "the nucleic acid disposed on beads" lacks antecedent basis; and
- In claims 92, 93, and 98, the phrase "said sequencing reagent" lacks antecedent basis.

Applicants have amended claims 84, 92, 93, and 98 for clarity. Specifically:

- In claim 84, the phrase "the nucleic acids disposed on beads" has been amended to read "nucleic acids in the wells;" and
- In claims 92, 93, and 98, the phrase "said sequencing reagent" has been amended to read "said pyrophosphate sequencing reagent."

These amendments are supported by the application as originally filed, and do not constitute new matter. It is believed that the amendments obviate this ground of rejection. Withdrawal of the rejection is respectfully requested.

Rejection under 35 U.S.C. § 103(a)

The Examiner has newly rejected claims 56-61, 64-68, 84-93, and 96-99 under 35 U.S.C. §103(a) as allegedly unpatentable over Chee et al. (U.S. Published Application No. 2003/1018867 ("Chee et al."); Office Action, page 7). The Examiner states that Chee et al. describe a substrate for pyrophosphate sequencing and an apparatus for the substrate. The Examiner notes that Chee et al. fail to teach: 1) imaging through a CCD (charged coupled device); 2) the diameter of the individual optic fiber or the depth of the cavitated fiber optic wafer; 3) various separation distances between the nucleic acids on the substrate/apparatus; and 4) an optical linkage between the polished end of the fiber optic wafer and a second fiber optic fiber (Office Action, page 9). However, the Examiner states that it would have been obvious to one of skill in the art to modify the teachings of Chee et al. to arrive at the invention as claimed (Office Action, page 9). Applicants respectfully traverse this rejection as follows.

1. ***Chee et al.* do not teach or suggest the coating limitations of the instant claims**

The Examiner states that Chee et al. expressly disclose the use of "photolithography" (paragraphs [0111] – [0114]) and thereby teach the subject matter of instant claims 90, 91, 96,

and 97 (Office Action, page 8). However, paragraphs [0111] – [0114] of Chee et al. do not contain the express teaching of instant claims 90, 91, 96, and 97, which recite:

> 90. The substrate of claim 56 wherein the cavitated fiber optic wafer is coated.
>
> 91. The substrate of claim 90 wherein the coating is selected from the group consisting of plastic, gold layers, organosilane reagents, photoreactive linkers, hydrophilic polymer gels and pluronic polymers.
>
> 96. The apparatus of claim 84 wherein the cavitated fiber optic wafer is coated.
>
> 97. The apparatus of claim 96 wherein the coating is selected from the group consisting of plastic, gold layers, organosilane reagents, photoreactive linkers, hydrophilic polymer gels and pluronic polymers (emphasis added).

Specifically, paragraphs [0111] – [0114] of Chee et al. do not teach or suggest cavitated fiber optic wafers which comprise a "coating," including a "plastic coating," or a coating of "gold layer," "organosilane reagents," "photoreactive linkers," "hydrophilic polymer gels," or "pluronic polymers." Paragraph [0114] of Chee et al. refers only to the use of "chemically modified sites" (paragraph [0114]). Thus, Chee et al. do not teach or suggest the subject matter of instant claims 90, 91, 96, and 97. MPEP 2142 states that for a *prima facie* case of obvious, a prior art reference must teach or suggest all the claim limitations. In addition, there must be a suggestion or motivation in the reference to modify the prior art to obtain the claimed invention. MPEP §2143.01 (see below). Here, neither requirement has been met. Therefore, it is respectfully asserted that instant claims 90, 91, 96, and 97 are not unpatentable over Chee et al. as cited by the Examiner.

2. *Chee et al.* do not teach or suggest the separation limitations of the instant claims

The Examiner states that Chee et al. expressly disclose "high density" arrays (paragraph [0105]) and thereby make obvious the subject matter of instant claims 59-61 (Office Action, page 8). However, paragraph [0105] of Chee et al. refers to arrays containing 40,000 beads/mm^2 to 1,000,000 beads/mm^2 (Office Action, pages 9-10). This corresponds to 40 beads/μm^2 to 1000 beads/μm^2, meaning that each bead would be separated by approximately 0.17 μm to 0.03 μm (see Declaration of Dr. Margulies under 37 C.F.R. §1.132, attached hereto).

In contrast, the instant claims recite:

59. The substrate of claim 58, wherein the fiber optic surface includes two or more nucleic acids separated by approximately 10 μm to approximately 200 μm.

60. The substrate of claim 58, wherein the fiber optic surface includes two or more nucleic acids separated by approximately 10 μm to approximately 150 μm.

61. The substrate of claim 58, wherein the fiber optic surface includes two or more nucleic acids separated by approximately 150 μm (emphasis added).

Accordingly, paragraph [0105] of Chee et al. does not teach or suggest fiber optic surfaces which comprise a "nucleic acids" that are separated by "approximately 10 μm to approximately 200 μm," "approximately 10 μm to approximately 150 μm," or "approximately 150 μm." Thus, Chee et al. paragraph [0105] does not teach or suggest the subject matter of instant claims 59-61. MPEP 2142 states that for a *prima facie* case of obvious, a prior art reference must teach or suggest all the claim limitations. In addition, there must be a suggestion or motivation in the reference to modify the prior art to obtain the claimed invention. MPEP §2143.01 (see below). Here, neither requirement has been met. Therefore, it is respectfully asserted that instant claims 59-61 are not unpatentable over Chee et al. as cited by the Examiner.

3. *Chee et al.* do not teach or suggest the compact wafer, detachable fiber, optic linkage, flow chamber, or fluid means of the instant claims

The Examiner states that Chee et al. fail to disclose the specific compact wafer, detachable fiber, and optical linkage recited in the instant claims (Office Action, page 9). Applicants add that Chee et al. also fail to disclose the specific flow chamber and fluid means recited in the claims. The Examiner states that that the compact wafer of the invention is merely ornamental and imparts no mechanical function (Office Action, page 10). Applicants respectfully emphasize that the recited compact wafer, detachable fiber, optical linkage, flow chamber, and fluid means impart notable *functional* advantages over the Chee et al. fiber, and other sequencing systems (see Dr. Margulies' Declaration, ¶¶ 5-14).

The substrate of instant claims includes a compact wafer between 0.5 mm and 5.0 mm in thickness (see, *inter alia*, page 36, lines 12-15; and page 36, line 30 to page 37, line 3). The instant application teaches that the wafer can be fit into a flow chamber and fluid means for delivering sequencing reagents and washes to the wafer surface (see, *inter alia*, Figures 2 and 3; page 4, lines 14-18; page 30, lines 15-19; page 33, line 20 to page 35, line 20; and Example 3 on page 53, line 28 to page 54, line 15). For sequence analysis, the underside of the claimed compact wafer in the flow chamber can be optically linked or directly contacted with a second optical fiber to allow image capture, for example, through a CCD system (see, *inter alia*, page 34, lines 13-18; and Figure 2). Alternatively, the underside of the compact wafer in the flow chamber can be placed in proximity to conventional optics mechanism, e.g., a high numerical aperture lens system to allow for image capture (see, *inter alia*, page 34, lines 19-23).

Due to the recited wafer size, flow chamber, fluid means, detachable fiber, and optical linkage the claimed substrate and apparatus yield significantly improved results compared to the system of Chee et al. (Dr. Margulies' Declaration, ¶ 10). The compact wafer of the invention is suitable for use flow chambers and fluid means, while the Chee et al. fibers are not (Dr. Margulies' Declaration, ¶¶ 10-11). Chee et al. (paragraph [0007], relying on WO 98/50782) are limited to the use of optical fibers extending several meters (see Dr. Margulies' Declaration, ¶ 11; WO 98/50782 page 14, lines 18-20).[1] The Chee et al. fibers are excessively long and therefore not compatible for use with flow chambers and fluid means (Dr. Margulies' Declaration, ¶ 11). Instead, Chee et al. report the use of a "dipping" mechanism, which involves inverting the long optical fiber and sequentially dipping the tip into individual cups filled with solutions of single nucleotides (Chee et al., ¶¶ [0192] – [0195], *inter alia*; Dr. Margulies' Declaration, ¶ 11).

The dipping mechanism of Chee et al. appears inoperable, or at best, unwieldy and inefficient for sequence analysis (Dr. Margulies' Declaration, ¶¶ 12-13). In contrast, the compact wafer of the invention can be placed into a flow chamber with fluid means to provide rapid and efficient delivery of sequencing reagents and washes (Dr. Margulies' Declaration, ¶¶ 13-14). Thus, the claimed wafer, detachable fiber, and optical linkage (and flow chamber with fluid means) are *not merely aspects of ornamentation*, but rather, highly advantageous *functional* features of Applicants' invention (Dr. Margulies' Declaration, ¶¶ 5 and 14).

[1] WO 98/50782 was submitted with Applicants' Information Disclosure Statement mailed February 1, 2002.

Thus, Chee et al. do not teach or suggest the subject matter of instant claims 56-61, 64-68, 84-93, and 96-99, including the recited compact wafer, detachable fiber, optical linkage, nucleic acid separation, flow chamber, and fluid means. MPEP 2142 states that for a *prima facie* case of obvious, a prior art reference must teach or suggest all the claim limitations. In addition, there must be a suggestion or motivation in the reference to modify the prior art to obtain the claimed invention. MPEP §2143.01 (see below). Here, neither has been shown by the Examiner. Moreover, the recited features of claims 56-61, 64-68, 84-93, and 96-99 are not merely "ornamentation." Rather, these features impart functional advantages over the system of Chee et al. and others (Dr. Margulies' Declaration, ¶¶ 5 and 14). Therefore, it is respectfully asserted that instant claims 56-61, 64-68, 84-93, and 96-99 are not unpatentable over Chee et al. as cited by the Examiner.

4. **The compact wafer, detachable fiber, and optic linkage of the instant claims are not equivalent to the fiber of *Chee et al.***

The Examiner states that because the claimed wafer *can* be attached to a second optical fiber bundle to transmit data to an imaging device, Applicants' invention is tantamount to the fiber of Chee et al. (Office Action, page 10). However, the instant application discloses that the second optical fiber bundle may be optically linked with the compact wafer (see, *inter alia*, page 34, lines 13-14; see also current claim 89). Thus, the light generated by the compact wafer can be transmitted to the exterior of the flow chamber to allow signal detection (see, *inter alia*, page 34, lines 13-18; and Figure 2).

Due to this detachable fiber and optical linkage, the compact wafer can be used in a flow chamber to allow rapid and efficient delivery of sequencing reagents and washes (Dr. Margulies' Declaration, ¶ 10). By comparison, Chee et al. are limited to the use of long, bulky fibers that necessitate a slow, inefficient, and wasteful "invert and dip" mechanism (Dr. Margulies' Declaration, ¶ 11). Therefore, the claimed wafer, detachable fiber, and optical linkage are not equivalent to the extended fibers of Chee et al. as cited by the Examiner.

5. *Chee et al.* do not teach or suggest the specific dimensions of the compact wafer as recited in the instant claims

The Examiner states that Chee et al. fail to disclose the specific diameter of the optical fibers of the compact wafer recited in the instant claims (Office Action, page 9). Applicants add that Chee et al. also fail to disclose the specific depths of the wells of the compact wafer recited in the claims. Applicants respectfully emphasize that these features impart notable *functional* advantages over the Chee et al. system (Dr. Margulies' Declaration, ¶¶ 15-30).

The substrate of instant claims includes a compact wafer between 0.5 mm and 5.0 mm in thickness (see, *inter alia*, page 36, lines 12-15; and page 36, line 30 to page 37, line 3). Additionally, the compact wafer includes optical fibers with a recited diameter of 3 mm to 100 mm (see, *inter alia*, page 36, line 15). The instant application teaches that this diameter is important to ensure that each light signal can be captured as a single pixel (see, *inter alia*, page 36, line 15 and 25-29). The compact wafer also includes wells with a recited depth of one-half to three times the diameter of the optical fibers (see, *inter alia*, page 37, lines 6-9). Relevant to this, the application expressly recognizes the problem of bead/sample loss during the sequencing reaction (see, *inter alia*, page 37, lines 6-9; page 39, lines 22-24; and Figure 4).

Due to the recited fiber diameter and well depth, the claimed substrate and apparatus yield significantly improved results compared to the system of Chee et al. (Dr. Margulies' Declaration, ¶ 18). Many of the fibers reported by Chee et al. are not useful for sequence analysis due to their excessively small size (Dr. Margulies' Declaration, ¶ 20). Chee et al. paragraph [0105] suggests the use of "high density" long optical fibers with diameters ranging from 0.17 μm to 0.03 μm (Dr. Margulies' Declaration, ¶ 19). These dimensions can be calculated from the range of optical fiber densities indicated by Chee et al. paragraph [0105], i.e., arrays containing 40,000 fibers/mm^2 to 1,000,000 fibers/mm^2 (Chee et al., ¶ [0105]; Dr. Margulies' Declaration, ¶ 19).

With diameters of approximately 0.17 μm to 0.03 μm, many of the optical fibers employed by Chee et al. produce sequencing systems that are completely or partly inoperable (see Dr. Margulies' Declaration, ¶ 20). For example, such systems would have problems in distinguishing light signals from each fiber and in depositing the beads in the wells (Dr. Margulies' Declaration, ¶ 20). In comparison to Chee et al. paragraph [0105], the compact wafer of the invention employs optical fibers 3 μm to 100 μm in diameter to provide for maximal

sample density while still allowing accurate signal detection and efficient bead delivery (Dr. Margulies' Declaration, ¶ 20).

The wells reported by Chee et al. do not appear to have any specified depth, suggesting that that Chee et al. have failed to recognize the importance of well depth in preventing sample loss. Figure 1 in Chee et al. shows beads and samples jutting out from their wells, which would seem to lead to significant sample loss during their "invert and dip" process (see Dr. Margulies' Declaration ¶ 21). In contrast, the compact wafer of the invention employs well depths of one-half to three times the diameter of the fiber to help minimize sample loss during preparation and analysis (see Dr. Margulies' Declaration, ¶ 21). The claimed dimensions of the fibers and wells therefore represent significant *functional advantages* over the system reported in Chee et al.

Thus, as cited by the Examiner, Chee et al. do not teach or suggest the subject matter of instant claims 56-61, 64-68, 84-93, and 96-99, including the including the recited compact wafer, fiber diameter, and well depth. MPEP 2142 states that for a *prima facie* case of obvious, a prior art reference must teach or suggest all the claim limitations. In addition, there must be a suggestion or motivation in the reference to modify the prior art to obtain the claimed invention. MPEP §2143.01 (see below). Here, neither has been shown by the Examiner. Moreover, the recited features of claims 56-61, 64-68, 84-93, and 96-99 represent functional advantages in view of Chee et al. (Dr. Margulies' Declaration, ¶ 21). Therefore, it is respectfully asserted that instant claims 56-61, 64-68, 84-93, and 96-99 are not unpatentable over Chee et al. as cited by the Examiner.

6. Conclusion: the instant claims are not obvious in view of *Chee et al.* as cited by the Examiner

First, the Examiner states that Chee et al. fail to teach a cavitated fiber comprising a depth between 0.5 mm and 5.0 mm, but alleges that the compact wafer is obvious since it imparts only ornamentation and no mechanical function (Office Action, pages 10-11). Yet, Applicants have demonstrated that the wafer element of the instant application imparts notable *functional* advantages over the Chee et al. system and others (Dr. Margulies' Declaration, ¶¶ 5-14 and 15-21)..

Applicants' recognize that, in some cases, the particular shape of a product is considered to lack patentable significance. MPEP § 2144.04; *Ex parte Hilton*, 148 USPQ 356 (Bd. App.

1965). However, the shape and size of the claimed wafer element, along with the detachable linkage, flow chamber and fluid means, fiber diameter, and well depth impart patentable significance, since these features results in substrates and apparatuses that are clearly distinct from and superior to other sequencing platforms, including that of Chee et al. (Dr. Margulies' Declaration, ¶¶ 10 and 18). *See* MPEP § 2144.04; *Ex parte Hilton*, 148 USPQ 356 (Bd. App. 1965).

The claimed invention, which includes the compact wafer and the other functional features, has been shown to outperform other sequencing devices and to allow whole-genome sequencing (Dr. Margulies' Declaration, ¶¶ 22-30). The Examiner is required to consider these inherent advantages, properties, utilities, and results flowing from the claimed invention, since each is a part of the invention as a whole. *In re Chupp*, 816 F.2d 643 (Fed. Cir. 1987); *Fromson v. Advance Offset Plate*, 755 F.2d 1549 (Fed. Cir. 1985); *In re Piasecki*, 745 F.2d 1468 (Fed. Cir. 1984); *Carl Schenck, AG v. Nortron Corp.*, 713 F.2d 782 (Fed. Cir. 1983); *In re Sernaker*, 702 F.2d 989 (Fed. Cir. 1983).

Second, the Examiner states that Chee et al. fail to teach numerous recited features of the claimed substrate and apparatus, including the compact wafer, fiber diameter, well depth, nucleic acid separation, optical linkage, and detachable optical fiber (Office Action, page 9). Applicants add that Chee et al. also fail to teach the recited flow chamber, fluid means, and well depth. However, the Examiner states that it is well within the purview of an ordinarily skilled artisan to make modifications to the system of Chee et al. and arrive at the invention as claimed (Office Action, page 9). It is further stated that modification of the Chee et al. system to obtain claimed invention would yield a reasonable expectation of success (Office Action, pages 9-10).

Yet, Applicants note that "although a prior art device may be capable of being modified to run the way the apparatus is claimed, there must be a suggestion or motivation in the reference to do so" in order to establish obviousness. MPEP §2143.01; *In re Mills*, 916 F.2d 680, 682 (Fed. Cir. 1990) (emphasis added). The fact that the claimed invention is alleged to be within the capabilities of one of ordinary skill in the art is not sufficient by itself to establish a *prima facie* case under Section 103. MPEP §2143.01; *In re Kotzab*, 217 F.3d 1365, 1371 (Fed. Cir. 2000).

Here, the Examiner has failed to provide any suggestion or motivation to modify the substrate of Chee et al. to obtain the *specifically claimed features* of the compact wafer, flow chamber, fluid means, fiber diameter, well depth, nucleic acid separation, detachable optical

fiber, and optical linkage of Applicants' invention. The Examiner relies on the "high density" arrays of Chee et al. paragraph [105] to allegedly show that the claimed separation distances between the nucleic acids is within the purview of a person in the art (Office Action, page 10). Yet, the high density arrays of Chee et al. correspond to separation distances of approximately 0.17 μm to 0.03 μm (Dr. Margulies' Declaration, ¶ 19). In contrast, the instant claims recite distances of "approximately 10 μm to approximately 200 μm," "approximately 10 μm to approximately 150 μm," or "approximately 150 μm." The Examiner provides no suggestion or motivation in Chee et al. paragraph [105] to employ the *specifically recited* distances. Moreover, the Examiner provides no suggestion or motivation in any part of Chee et al. to arrive at the specifically recited features of the compact wafer, flow chamber, fluid means, fiber diameter, well depth, optical linkage, and detachable optical fiber of the instant claims.

Thus, it is respectfully submitted that a *prima facie* case of obviousness has not been established against the instant claims. In view of all of the above, Applicants respectfully submit that claims 56-61, 64-68, 84-93, and 96-99 (as well as new claim 100) are not unpatentable over Chee et al. as cited by the Examiner. Reconsideration of the pending claims is respectfully requested.

Declaration under 37 C.F.R. §1.132 from Marcel Margulies filed October 9, 2003

In the Office Action, the Examiner failed to considered the Rule 132 Declaration of Dr. Marcel Margulies, which was filed in the U.S. Patent and Trademark Office on October 9, 2003 (see stamped Transmittal Letter and Declaration (page 1); Exhibit 14). Applicants' have resubmitted the contents of this Declaration in the new Rule 132 Declaration submitted herewith. Entry and consideration of the newly filed Rule 132 Declaration of Dr. Margulies is respectfully requested.

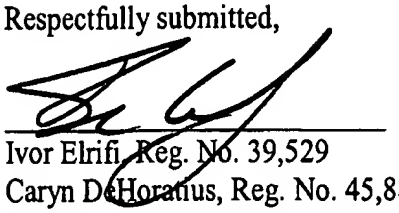
CONCLUSION

Applicants believe that the claims as amended are patentable and a prompt allowance is respectfully requested. If further discussion of this case is deemed helpful, the Examiner is encouraged to contact the undersigned at the telephone number provided below, and is assured of full cooperation in progressing the instant claims to allowance. While Applicants believe that no additional fees are required, the Commissioner is authorized to charge or credit the undersigned

Deposit Account No. 50-0311, Reference No. 21465-501 CIP2, Customer No. 35437, for any additional fees needed.

Dated: April 23, 2004

Respectfully submitted,



Ivor Elrifi, Reg. No. 39,529
Caryn DeHoratius, Reg. No. 45,881
MINTZ, LEVIN, COHN, FERRIS,
GLOVSKY and POPEO, P.C.
The Chrysler Center
666 Third Avenue, 24th Floor
New York, NY 10017
Tel: (212) 935-3000
Fax: (212) 983-3115

Express Mail Label No.: EV287336791US
Date of Deposit: April 23, 2004

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS:	Jonathan M. Rothberg, et al.		
ASSIGNEE:	CuraGen Corporation		
SERIAL NUMBER:	09/814,338	EXAMINER:	Young J. Kim
FILING DATE:	March 21, 2001	ART UNIT:	1637
FOR:	METHOD OF SEQUENCING A NUCLEIC ACID		

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

DECLARATION OF MARCEL MARGULIES, Ph.D. UNDER 37 C.F.R. §1.132

I, MARCEL MARGULIES, declare and state that:

1. I am Vice President of Engineering, at 454® Life Sciences, the exclusive licensee of this application. My previous employment includes Director of New Technology Research at Perkin-Elmer's Instrument Division in Norwalk, CT, and Associate Director of the Hubble Space Telescope project.

2. I earned my B.Sc. in Engineering from the Free University of Brussels, in Belgium, and a Ph.D. in theoretical physics from Columbia University.

3. I have reviewed the instant application and the November 6, 2003 Office Action in this case.

4. Based on information and belief, it is my opinion that the claimed sequencing apparatus and substrate of the instant application are vastly superior to other sequencing systems, including the system reported by Chee et al. in Published Application No. U.S. 2003/0108867 ("Chee et al."), as cited by the Examiner.

The superior performance of the claimed invention can be attributed to the functional features of the apparatus and substrate, which include: 1) compact wafers; 2) attachable optical fibers; 3) flow chamber and fluid means; and 4) specific fiber and well sizes. As a result of these features (and others), the claimed invention provides the first massively parallel, solid-phase, whole-genome sequencing platform that can be scaled for viral, bacterial, and even human genomes.

Functional features of the claimed substrate and apparatus:

compact wafers can be placed in flow chambers for efficient fluid exchange

5. As recited in the claims, the compact wafers of the invention allow *placement into flow chambers*, which utilize an *efficient fluid exchange system* and thereby provide significantly faster sequence analysis. The advantages of the claimed apparatus and substrate are fully disclosed in the instant application, as filed.

6. The instant application teaches that the claimed apparatus and substrate include a compact wafer formed from a bundle of optical fibers, cut and polished to a thickness of 0.5 mm to 5 mm. This teaching is found, *inter alia*, in the originally filed application on p. 36, l. 12-15; and p. 36, l. 30 to p. 37, l. 3.

7. The instant application additionally teaches that the claimed substrate can be used with a flow chamber and fluid means for delivering sequencing reagents and washes to the wafer surface. This teaching is found, *inter alia*, in the originally filed application in Figs. 2 and 3; on p. 4, l. 14-18; p. 30, l. 15-19; p. 33, l. 20 to p. 35, l. 20; and in Exmp. 3 on p. 53, l. 28 to p. 54, l. 15.

8. The instant application further teaches that the underside of the compact wafer in the flow chamber can be optically linked or directly contacted with a fiber optic bundle to allow image capture, for example, through a CCD system. This teaching is found, *inter alia*, in the originally filed application on p. 34, l. 13-18; and Fig. 2.

9. The instant application teaches also that the underside of the compact wafer in the flow chamber can be placed in proximity to conventional optics mechanism, e.g., a high numerical aperture lens system to allow for image capture. This teaching is found, *inter alia*, in the originally filed application on p. 34, l. 19-23.

10. The claimed wafer for use in the flow chamber allows for much faster sequence analysis. The claimed apparatus and substrate thereby yield significantly improved results, which are not obtained with other sequencing systems such as that reported by Chee et al.

11. As cited by the Examiner, Chee et al. does not specify the length of the optic fibers for their sequencing system. Instead, Chee et al. relies on WO 98/50782 (see Chee et al., ¶ [0007]), which reports the use of optic fibers that are several meters long (Ex. 1). These long and bulky fibers cannot readily fit into a flow chamber. Instead, Chee et al. reports methods of inverting the long optic fiber and sequentially dipping the tip into individual cups filled with solutions of single nucleotides (Chee et al., ¶¶ [0192] – [0195], *inter alia*). This awkward dipping process is necessitated by the long, bulky fiber employed in Chee et al.

12. The dipping method of Chee et al. is predicted to be completely or partly inoperable. First, the continuous plunging of an optic fiber tip into nucleotide solutions would tend to dislodge any beads from the wells. Second, for DNA attached directly to wells, dipping would be ineffective in delivering the nucleotide solution to the wells due to the counteraction of air pressure. This phenomenon is generally observed for inverted cups or glasses placed into reservoirs of water, and provides the basis for the oceanographic apparatus known as the diving bell. An illustration provided as an aid in understanding is included as Ex. 2 (adapted from http://home.earthlink.net/~dmocarski/chapters/chapter7 /main.htm).

13. Even assuming, *arguendo*, that the dipping method of Chee et al. were marginally effective in delivering reagents and preserving samples for sequencing, the method would be extremely slow and inefficient compared to the claimed invention. The long fiber optic tip of Chee et al. would need to be inverted and dipped in and out of at least four cups (e.g., A, T, C, G), or perhaps more than eight cups (e.g., A, first wash, T, second wash, C, third wash, G, fourth wash) to determine only one nucleotide of sequence. In contrast, compact wafer of the invention can be placed into a flow chamber to allow for rapid and efficient delivery of sequencing reagents and washes to the compact wafer. An illustration provided as an aid in understanding is included as Ex. 3.

14. Because the long optic fibers of Chee et al. cannot readily fit into a flow chamber, the sequencing reactions are performed using an unwieldy dipping process. This results in significant delays and increased sample losses. By comparison, the claimed wafer is fitted into a flow chamber to allow streamlined processing and sequence analysis. This is a *significant functional advantage* over the system reported in Chee et al.

Functional features of the claimed substrate and apparatus:

compact wafers have well sizes to maximize signal capture and minimize sample loss

15. As recited in the claims, the compact wafers of the invention include *optimally sized fibers and wells* that allow *maximal signal capture and minimal sample loss* and thereby provide significantly improved sequence analysis. The advantages of the claimed apparatus and substrate are fully disclosed in the instant application, as filed.

16. The instant application teaches that the claimed apparatus and substrate include a compact wafer that includes optic fibers with a diameter of 3 μm to 100 μm. The application teaches that this diameter is important to ensure that each light signal can be captured as a single pixel. This teaching is found, *inter alia*, in the originally filed application on p. 36, l. 15 and 25-29.

17. The instant application teaches that the claimed apparatus and substrate include a compact wafer that includes wells with a depth of one-half to three times the diameter of the optic fibers. The application expressly recognizes the problem of bead/sample loss during the sequencing reaction. This teaching is found, *inter alia*, in the originally filed application on p. 37, l. 6-9; p. 39, l. 22-24; and Fig. 4.

18. The claimed fiber diameter and well depth of the compact wafer allow for much more effective sequence analysis. Accordingly, the claimed apparatus and substrate yield significantly improved results, which are not obtained with other sequencing systems, including the system reported by Chee et al.

19. As cited by the Examiner, Chee et al. ¶ [0105] apparently reports the use of long optic fibers with diameters ranging from approximately 0.17 μm to 0.03 μm. These diameters can be calculated from "high density" arrays indicated by Chee et al. ¶ [0105], i.e., arrays containing 40,000 fibers/mm^2 to 1,000,000 fibers/mm^2 (Office Action, quoting Chee et al. on pages 9-10). Chee et al. appears to be silent as to the specific well depths employed with the optic fibers.

20. With diameters of approximately 0.17 μm to 0.03 μm, many of the optic fibers employed by Chee et al. would produce sequencing systems that are completely or partly inoperable. Optic fibers having a such small diameters would require bead and well sizes less than 0.17 μm to 0.03 μm in diameter. Such systems would be predicted to have a myriad of problems, including difficulties in distinguishing light signals from each fiber and in depositing the beads in the wells. By comparison to Chee et al. ¶ [0105], the compact wafer of the invention employs optic fibers 3 μm to 100 μm in diameter to provide for maximal sample density while still allowing accurate signal detection and efficient bead delivery.

21. Chee et al. do not appear to specify well depths for use with the optic fibers, and evidently fail to recognize the importance of well depth in preventing sample loss.

In fact, Fig. 1 in Chee et al. shows beads and samples jutting out from their wells. This configuration would likely lead to significant sample loss during the "invert and dip" process reported by Chee et al. In contrast, the compact wafer of the invention employs well depths of one-half to three times the diameter of the fiber, which are important in minimizing sample loss during preparation and analysis. The optimally sized fibers and wells (¶¶ 16 and 17, above) therefore represent *significant functional advantages* over the system reported in Chee et al.

Superior function of the claimed substrate and apparatus:
massively parallel analysis of viral and human genomic sequences

22. As a result of these highly advantageous, functional features (¶¶ 14, 20, and 21, above), and other important aspects, the substrate and apparatus claimed in the instant application are the first to allow rapid massively parallel sequencing for whole genomes.

23. Traditional methods for genome sequencing have been slow, expensive, laborious, and industrial-scale, since they involve individually preparing and sequencing DNA fragments of the genome. The Human Genome Project, for example, required approximately 12 years, $2.7 billion dollars, and 60 million samples to complete.

24. In contrast, the substrate and apparatus claimed in the instant application provide a massively parallel, scalable platform that dramatically reduces the time, cost, sample preparation, and space required for genome sequencing. Instead of individually preparing and sequencing each sample, the claimed substrate and apparatus allow parallel sequencing of thousands (or hundreds of thousands) of samples.

25. Recently, the claimed substrate and apparatus were used to sequence the entire adenovirus genome (approximately 30,000 base pairs) contained on an expression vector in less than one day (see NY Times article, Ex. 4). The entire sequencing

process from sample preparation to data analysis was accomplished in less than one day, and provided over 99% genome coverage. The resulting adenovirus sequence was published in GenBank under Accession Nos. AY370909, AY370910, and AY370911 (Ex. 5).

26. In further experiments, the apparatus of the instant application was used to sequence a segment human chromosome 12 (approximately 170,000 base pairs) contained on an artificial chromosome vector (Ex. 6). With the apparatus, a one-day sequencing run produced sufficient shotgun sequence coverage of the chromosome 12 clone (Ex. 6, p. 6). A single sequencing run obtained 85% genome coverage and 98% consensus accuracy (Ex. 6, p. 3). These results were presented at the15th Annual Genome Sequencing and Analysis Conference, held on September 21-24, 2003 (Ex. 6, p. 1).

27. To generate this sequence information described in ¶¶ 25 and 26 (above), preferred commercial embodiments of the claimed substrate and apparatus were fabricated. In these preferred embodiments, the claimed substrates (termed "PicoTiter Plates") were formed from cavitated fiber optic wafers formed from a fused bundle of a plurality of individual optical fibers as taught and claimed by the instant application.

28. Specifically, PicoTiter Plates were made acid etching the top surface of fiber optic wafers to form wells with diameters between 39 and 44 μm, as currently claimed. The fiber optic wafer exhibited a thickness of about 2.0 mm, also as currently claimed. In addition, the wells on PicoTiter Plates were fabricated with depths ranging from 26 to 76 μm (i.e., from between one half the diameter of an individual optical fiber and three times the diameter of an individual optical fiber, as recited in the pending claims). Finally, the wells were loaded with nucleic acid template and beads with pyrophosphate sequencing reagents attached thereto, as recited in the pending claims. Sequencing by synthesis was then performed as

described in the specification, and using the claimed apparatus to flow sequencing reagents over the PicoTiter Plate.

29. The substrate and apparatus claimed in the instant application therefore fulfill a long-felt but unmet need for rapid, whole-genome analysis of viral and bacterial pathogens (e.g., ¶ 25, above). Such analysis is critical for biodefense, drug discovery, and the identification of emerging pathogens. More than this, the claimed apparatus solves the long-standing problems with analysis of large genomes, such as human genomes (e.g., ¶ 26, above). Solutions for large-genome sequencing are vital for drug development, early diagnosis, and faster clinical interventions.

30. For these reasons, in my opinion, the claimed substrate and apparatus represent a significant advancement in the field as the first massively parallel, solid-phase, whole-genome sequencing platform that can be scaled for the smallest to the largest genomes.

Conclusion

31. Therefore, based on information and belief, and all of the foregoing, it is my opinion the claimed sequencing apparatus and substrate substantially outperform the sequencing platforms used by Chee et al. and others. This is due to the functionally superior features of the claimed invention, which include compact fiber optic wafers, detachable fiber optic bundles, flow chambers and fluid means, and specifically sized fibers and wells. All of these features, and the other aspects of the invention, work together to achieve significantly faster results compared to other sequencing systems.

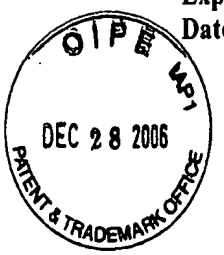
32. I declare that all statements made herein of our own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18

U.S.C. § 1001 and that willful false statements may jeopardize the validity of this application and any patent issuing therefrom.

Dated: 4/20/04

Signed: _____
MARCEL MARGULIES

Express Mail Label No.: EV781047676US
Date of Deposit: December 28, 2006



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICANTS:	Jonathan M. Rothberg, et al.		
ASSIGNEE:	454 Corporation		
SERIAL NUMBER:	09/814,338	EXAMINER:	Young J. Kim
FILING DATE:	March 21, 2001	ART UNIT:	1637
FOR:	METHOD OF SEQUENCING A NUCLEIC ACID		

Commissioner for Patents
P. O. Box 1450
Alexandria, VA 22313-1450

DECLARATION OF MARCEL MARGULIES UNDER 37 C.F.R. §1.132

I, MARCEL MARGULIES, declare and state that:

1. I am Vice President of Engineering, at 454 Life Sciences, the exclusive licensee of this application. My previous employment includes Director of New Technology Research at Perkin-Elmer's Instrument Division in Norwalk, CT, and Associate Director of the Hubble Space Telescope project.

2. I earned my B.Sc. in Engineering from the Free University of Brussels, in Belgium and a Ph.D. in theoretical physics from Columbia University.

3. I have read the specification and claims of the above-referenced patent application and the December 29, 2005 Office Action. I understand that there is a sole remaining rejection; the Examiner has rejected the claims as obvious over Chee et al., U.S. 2003/01808867 ("Chee") in view of Krull et al, WO 98/58079, ("Krull"). I provide this declaration to explain why the claims are not obvious over the cited art.

EXHIBIT E

4. It is my opinion that the claimed invention provides a cavitated fiber optic wafer for massively parallel, scalable platform that dramatically reduces the time, cost, sample preparation, and space required for genome sequencing. The substrate wafer claimed here therefore fulfill a long-felt but unmet need for large scale sequencing, and rapid whole-genome analysis.

5. In this patent application, the claims are directed to a substrate comprising a cavitated fiber optic wafer formed from a fused bundle of a plurality of individual optical fibers, each individual fiber having specified dimensions (or to an apparatus having such a substrate). Specifically, each of the claims require that each individual optical fiber has a diameter between 3 and 100 μM, the thickness of the wafer (i.e., length of the optic fiber) between the top surface and the bottom surface is between 0.5 mm and 5.0 mm and the depth of each well ranges from between one half the diameter of an individual optical fiber and three times the diameter of an individual optical fiber. These specific parameters of the claimed cavitated fiber optic wafers are not chosen arbitrarily.

6. I have provided below data demonstrating that the claimed parameters of the claimed cavitated wafers perform differently from long fiber optical bundles disclosed in Chee and/or Krull. First, the claimed wafer thickness is important for light transmission properties of the claimed wafer. Second, the claimed well diameter and well depth parameters are important to the diffusion properties of claimed substrate.

7. With respect to the light transmission properties of the claimed wafer, I provide below data showing that light transmission through the individual fibers varies significantly as a function of the length of the fiber. As summarized in Figure 1, the transmission for a fiber optic measuring 300 mm in length is less than 18%. In contrast, nearly 100% transmission is achieved for an instantly claimed fiber optic measuring in length between 0.5mm and 5.0mm (I note that the data plotted does not take into account the fact that the actual mean path in any one fiber is

longer than the linear fiber length -- which would exacerbate the differences between the short wafer thicknesses claimed and the long fiber bundles in the prior art).

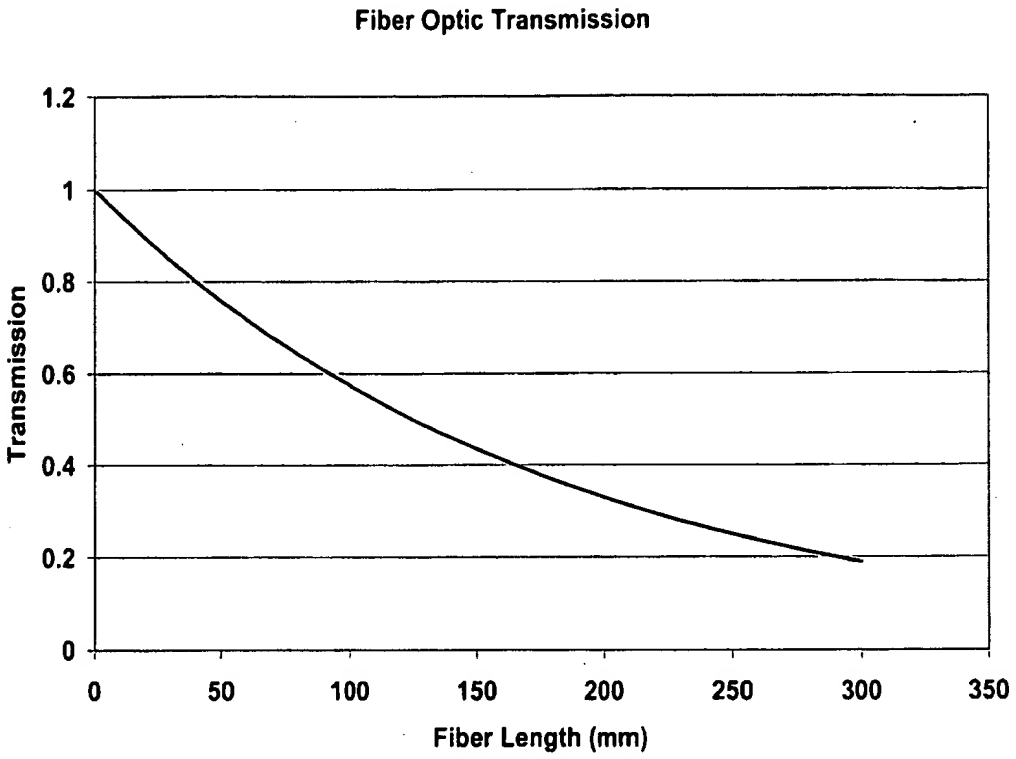


Figure 1.

8. Furthermore, the claimed parameters for well diameter (between 3 and 100μM) and well depth ranging from between one half diameter of an individual optical fiber and three times the diameter of an individual optical fiber are not arbitrarily chosen parameters. Well depth is selected on the basis of a number of competing requirements in a nucleic acid sequencing application: (1) wells need to be deep enough for DNA-carrying beads to remain in the wells in the presence of convective transport past the wells; (2) the wells must be sufficiently deep to provide adequate isolation against diffusion of by-products from a well in which incorporation is taking place to a well where no incorporation is occurring; (3) they must be shallow enough to allow rapid diffusion of nucleotides into the wells and rapid washing out of remaining nucleotides at the end of each flow cycle to enable high

sequencing throughput and reduced reagent use; and (4) they must not be so deep that it would be easy for more than one bead to fit in a well.

9. To assess the sensitivity of this system to reaction by-products diffusing from one well into a neighboring one, a simplified one-dimensional model of interwell diffusion behavior was developed. We have found that at a well-to-well distance of 50 μm, diffusion of ATP produced during a pyrophosphate-based sequencing reaction, will induce a background signal on the order of 10% or less in an immediately neighboring well. The timescale for diffusion into and out of the wells is on the order of 10 s in this configuration.

10. We further created a one-dimensional model of the claimed fiberoptic cavitated wafer (i.e. modeled a linear array of wells) in which the wells are represented as lumped chemical reactors that produce pyrophosphate and ATP during the sequencing reaction. Within each well the generation of reaction by-products can be modeled by a set of coupled kinetic equations. Numerical solution of this set of equations is illustrated by Figure 2.

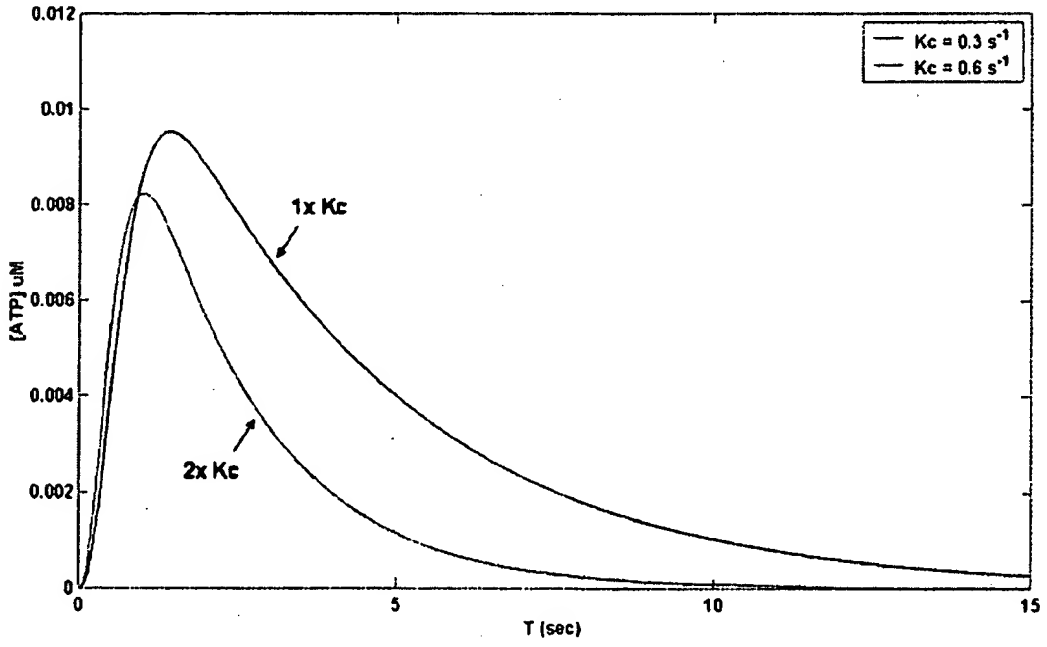


Figure 2.

11. As summarized in Figure 2, the simulation shows impact of well depth on signal generation and decay kinetics for a deeper (red) and shallower (green) well. The shallower well shows lower amount of signal (equal to the area under the curve) and faster signal decay (ATP leaves the well faster). Faster decay allows for reduction of flow cycle time interval, provided the signal generated is adequate. The deeper well results in a stronger signal (and slower signal decay) with the added caveat that too much signal, over too long a period of time, will result in deleterious optical bleed into neighboring wells. The effects of the above described observations on the experimental results of actual sequencing runs are summarized in Table 1.

Table 1.

Well Depth	40µm	35µm	30µm
Signal Per Base	1117 ± 267	1006 ± 226	787 ± 184
Diffusive signal bleed to down-stream well	13%	13%	17%
Doubly occupied wells	0.48%	0.27%	0.27%

12. In these experiments, three PTP's with different well depths (30 µm, 35 µm and 40 µm; all having a fiber optic diameter of approximately 35 µm) were used to sequence reference test fragments. As summarized in Table 1, the 40 µm deep wells show lower *chemical* cross-talk (signal bleed to downstream well). In contrast, the fraction of multiply occupied wells is twice as high as for 35 µm or 30 µm deep wells. The 30 µm deep wells have a low fraction of multiply occupied

wells, but higher chemical cross-talk, and lower signal per base. Finally, the 35 μm wells have lower chemical cross-talk, a lower fraction of multiply occupied wells and adequate signal/base.

13. Table 2 summarizes actual sequencing results that bear out these findings.

Table 2

Well Depth	Key Pass	Read Error	Matching key pass reads over 100 bases		
			100%	98%	95%
40	17440	0.46%	81.84%	94.56%	96.78%
35	17928	0.40%	85.99%	95.29%	97.41%
30	17577	0.48%	79.19%	94.21%	97.21%

As summarized in Table 2, the read error is lowest for the 35 μm deep wells and the fraction of reads perfectly sequenced (matching at 100% over 100 bases) is highest for those wells. The results summarized in Table 2 unequivocally show the importance of the claimed well diameter and well depth parameters to achieve desired sequencing results. Neither Chee nor Krull teach or suggest the use of wells with depth ranging from between one half the diameter of an individual optical fiber and three times the diameter of an individual optical fiber.

14. The Examiner cites Chee for obviousness but concedes that Chee does not disclose the use of a cavitated fiber optic bundle as a wafer with thickness (i.e. length of the optic fiber) between 0.5mm and 5mm. The Examiner points to Krull to provide disclosure of a wafer (the Office Action cites to Krull, p. 13, lines 16-19). However, as discussed with the Examiner, Krull's optical "wafers" are completely unlike the optical fiber bundles claimed here (or the long fiber bundles referred to in Chee). Krull actively seeks to use the entire circumference of a single fiber for light transmission -- and to this end actually dissolves the external cladding from optical fiber pieces (see, e.g., Krull, Ex. 1, p. 43, lines 7-10; Ex. 2, lines

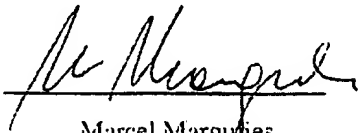
20-23: Figure 4(b)). In contrast, in the instant cavitated fiber optic bundle wafer, light transmission through the fiber is desired; light transmission between fibers (prevented by the cladding) is to be minimized or avoided. For this reason, Krull teaches directly away from the instantly claimed wafer and for this reason cannot be combined with Chee.

15. As demonstrated above, Chee is silent as to the specifically recited dimensions of the cavitated fiberoptic substrate claimed here. And, the data presented here demonstrates that the claimed dimensional parameters were not arbitrarily chosen, but are important in determining the light transmission and diffusional characteristics of the claimed wafer.

16. In summary, neither Chee nor Krull, alone or in combination (and they cannot be combined) teach or suggest all the claimed elements, and there is no suggestion to modify Chee with Krull to obtain the claimed invention. For the reasons outlined here, I believe that Chee and Krull cannot make obvious the subject matter of the present claims.

17. I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. § 1001 and that willful false statements may jeopardize the validity of this application and any patent issuing therefrom.

Dated: 12/22/06

Signed: 

Marcel Margulies